PROTEOMIC CHARACTERIZATION OF *FUSOBACTERIUM NUCLEATUM* OUTER MEMBRANE VESICLES

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

JINJING LIU

AUGUEST 2018

© 2018 JINJING LIU

ABSTRACT

Fusobacterium nucleatum is a Gram-negative anaerobic bacterium prevalent in the human oral cavity. This oral commensal bacterium has been shown as the main contributor for many dental diseases, including periodontitis and gingivitis. Recently, evidence has verifying the linkage between *F. nucleatum* and carcinogenesis of colorectal cancer (CRC). Outer membrane vesicles (OMVs) are naturally occurring metabolic byproducts secreted by Gram-negative bacteria, and contain highly antigenic LPS and outer membrane protein residues. Due to its antigen containing and non-self-replicated properties, OMVs are emerging as a powerful platform for vaccine design and delivery. We speculated the OMVs from *F. nucleatum* might contain antigens able to elicit effective immune response. In this study, we identified 98 proteins from purified OMVs of *F. nucleatum* through proteomic studies. *In silica* analysis indicates the non-replicative OMVs of *F. nucleatum* contain a variety of antigenic virulence factors and may play a critical role in vaccine design.

BIOGRAPHICAL SKETCH

Jinjing Liu is a second-year Master's student at Cornell University. She joined Dr. Yung-Fu Chang's lab at Cornell College of Veterinary Medicine in August 2016 and participated the project *Anti-Fusobacterium nucleatum Vaccine for Colorectal Cancer*. Before joining Cornell, Jinjing worked in the Institute of Zoology in Chinese Academy of Science and Henan Institute of Science and Technology from July 2015 to August 2016. Jinjing was involved in several projects including: *Infectious Bronchitis Virus* and *Porcine Immune Organ Development with Probiotic Feed*. Jinjing obtained her Bachelor of Science degree from University of Washington in June 2015, majoring in Molecular Cellular and Developmental Biology.

During her time at University of Washington, Jinjing volunteered in multiple organizations and projects contributing to education and science. She was involved in a Pipeline Project and worked as a volunteering teaching assistant at Garfield High School. She also facilitated the Foundation of International Understanding Through Students (FIUTS) and worked with a local elementary school in a combined cultural exchange and science education program. Additionally, she participated in an e-Health project led by the Biomedical and Health Information Department at UW to investigate the seniors' attitude toward companion robot pets in retirement communities through focus group interviews. These experiences underscored her belief in promoting public health and science education.

ACKNOWLEDGMENTS

I would first like to express my sincere appreciation for my advisor Prof. Yung-Fu Chang, whose knowledge and patience guided and supported my research and work here at Cornell. I would also like to thank my committee members: Prof. Xingen Lei and Prof. Hussni Omar Mohammed for their encouragement and insightful comments, which served to widen my perspective in my area of study. My sincere thanks also go to Dr. Ching-Lin Hsieh, Dr. Sheng Zhang, John L. Grazul and all members in Chang's lab, whose support proved to be vital for my project. In particular, I would like to thank my parents for continuous support and encourage me to complete this phase of my career.

This study was funded by HHSN2612015000391 and 1S10 OD017992-01 from the National Institutes of Health (NIH), USA, and Chinese Scholarship Consult (CSC), China.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE

INTRODUCTION AND LITERATURE REVIEW

The gastrointestinal (GI) track has the largest internal surface area in the human body and harbors several trillion microorganisms [1]. As the main segments of GI track, intestines are heavily colonized by complex and dynamic microbial communities. In the past, the role of microbiota was neglected, while in recent years this view has dramatically changed as researchers have uncovered the fact that microorganisms play a crucial role in host homeostasis and disease, ranging from obesity and infectious disease to neurodegenerative diseases [2-3]. Five major phyla A*ctinobacteria, Bacteroidetes, Firmicutes, Proteobacteria* and *Verrucomicrobia* make up the bulk of intestinal microbiota, which is highly diverse and abundant [4]. Studies have long established causation or association of main gut bacteria including *Clostridium difficile* (a key member of *Firmicutes*), *Escherichia coli* (a key member of *Proteobacteria*), *Bacteroides fragilis* (a key member of *Bacteroidetes*) and *Fusobacterium nucleatum* with dysbiosis and multiple human diseases such as inflammatory bowel disease (IBD) and colorectal carcinoma (CRC) [5-7]. Recently, the composition of intestinal flora has proven strongly correlated with diet and hygiene; investigations linking diseases progression with intestinal microbial communities are still ongoing.

1. 1 *Fusobacterium nucleatum*

1.1.1 Biology and Taxonomy of *Fusobacterium nucleatum*

Fusobacterium nucleatum is a Gram-negative non-sporefroming and obligate anaerobic (up to 6% oxygen) bacterium with fusiform rod shape; it is a key component of plaque and a common resident of human oral cavity [8-9]. It is a heterogeneous species and one of the most abundant organisms identified among endodontic infections [8, 10]. The name of *Fusobacterium* came from *fusus* and *bacterion*, which means spindle and rod [8].

In 1693, Antony van Leeuwenhock first detected the fusiform organisms from dental plaque, and later in 1894 and 1896, the bacteria were found in gingivitis samples and angina by Plaut and Vincent, respectively [11]. The fusiform bacteria were first isolated from the culture in 1898 by Veillon and Zuber and named *Bacillus fusiformis* [11]. In 1907 Leiner first recognized different culture types, and in 1913 Krumwied and Pratt first classified the fifteen strains into two groups based on their fermentability of sucrose [12]. In 1922, Knorr proposed the genus *Fusobacterium* and named three subspecies: *nuceatum, polymorphum,* and *plauti-vincentii* [8, 11-12]. In 1990 and 1992, Dzink et.al and Gharbia et.al grouped *Fusobacterium* into five subspecies: *nucleatum*, *polymorphum*, *fusiforme, animalis*, and *vicentii* [8, 11, 13]. Up to 2002, there were totally six subspecies been recognized including *F. canifelium,* which is isolated from cat and dog bite wounds in human [14]. Accumulating evidence has indicated among all subspecies *F. nucleatum* are frequently associated with diseases, which including periodontitis, adverse pregnancy outcomes, Lemierre's syndrome, inflammatory bowel disease, and colorectal cancer [15].

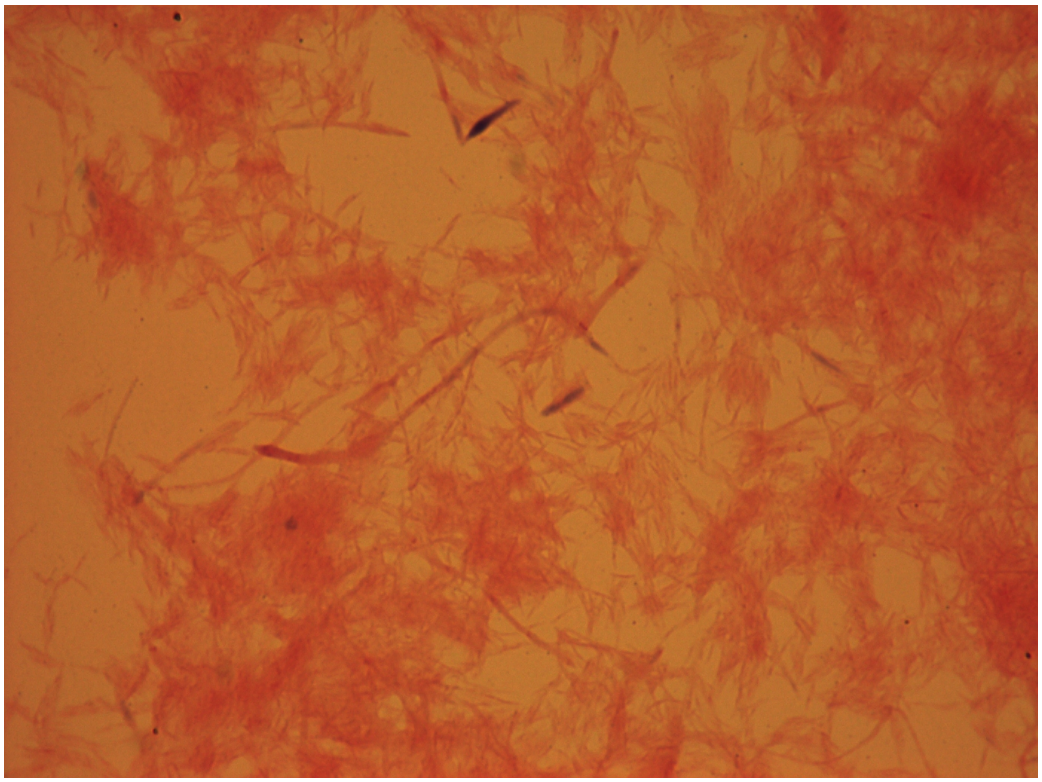


Figure 1 Microscope image of *F. nucleatum*

Gram stain of *F. nucleatum* appears red. *F. nucleatum* is an anaerobic bacterium with slender and spindle shape under the microscope.

1.1.2 Pathogenesis

F. nucleatum was first identified among plaque and found associated with dental disease. With the advancement and magnificent of experimental technology *F. nucleatum* was found prevalent among humans and has been implicated not only various oral diseases but GI disorders and other chronic infections such as Lemierre's syndrome and Alzheimer's disease as well [15-17]. As an oral commensal bacterium, *F. nucleatum* pervades the host mainly by colonization and dissemination to invade the epithelial cells [15]. Several virulence factors encoded by *F. nucleatum* such as FadA, Fap2, and RadD were shown able to adhere and interact with host cells [18-20].

FadA is the best-characterized adhesion molecule and uniquely encoded by *F. nucleatum*. It has an alpha-helical hairpin structure [20]. There are two forms of FadA: pre-FadA (nonsecreted from) containing 129 amino acids and mature FadA (secreted form) containing 111 amino acids [20]. A mixture of pre-FadA and mature FadA results in attachment to intestinal epithelium; however, preparations of mature FadA show no binding [20]. In addition to adhesion, induction of host immune response is another key virulence mechanism of *F. nucleatum*. To stimulate colorectal tumor progression, FadA binds to the E-cadherin receptor and actives β-catenin, resulting in upregulation of transcriptional factors, Wnt genes, inflammatory gene and oncogenes [21].

Another CRC associated virulence factor is lectin Fap2, a galactose adhesion protein that recognizes Gal-GalNAc, facilitates localization and colonization of *F. nucleatum* on CRC cell [22]. Additionally, Fap2 suppresses T cells and natural killer (NK) cells activities by directly interacts with inhibitory receptor TIGIT to evade the immune system attack [23].

1.1.3 *F. nucleatum* in Periodontal Disease

Porphyromonas gingivalis, Streptococcus sangius, Streptococcus gordonii, Tannerella forsythia, Treponema denticola, Actinobacillus actinomycetemcomitans, and *F. nucleatum* were primary etiologic agents of periodontitis [8, 24-25]. *F. nucleatum* is abundant in human oral cavity and is present in much higher levels in patients with periodontitis and gingivitis [25-28]. Samples collected from patients with gingivitis, chorionic periodontitis, and healthy groups were analyzed by real-time PCR and the experimental results displayed that the occurrence of *F. nucleatum* was positively correlated with the progression of the disease [25-26]. The elevation of IgG and IgA response levels were also detected in chronic periodontal patients [29]. It has been implied the combination of *F. nucleatum* and *P. gingivalis* or other bacterium was more pathogenic than *F. nucleatum* only infection [8, 30]. *F. nucleatum* recognized lactose, galactose and other sugars of *P. gingivalis* capsid and conjugated with these though carbohydrate-hydrate interaction [31-32]. The corncob receptor of *F. nucleatum* interacts with *P. gingivalis* to form coaggregation units, playing an essential role in attachment and plaque formation [24, 33-35]. In addition, *F. nucleatum* had the ability to reduce oxygen in the microenvironment thus facilitating *P. gingivalis* growth to promote periodontitis [36].

1.1.4 *F. nucleatum* in Colorectal Carcinoma

The primary role of *F. nucleatum* in the aetiology of CRC was revealed recently and has emerged as a focus of study. These studies have shown that presence of *F. nucleatum*, previously considered only as an intestinal commensal, has been linked with CRC progression and metastasis as an essential risk factor [37-38]. *F. nucleatum* bacterial loads from colorectal tumor tissue samples were

significantly higher compared with adjacent normal tissues tested using qPCR [37, 39]. This indicated that *F. nucleatum* could be used as a putative biomarker for CRC diagnoses. Recent studies have also indicated that *F. nucleatum* not only participated colorectal carcinogenesis but was interrelated with CRC outcomes as well [38]. Flanagan et al. discovered low *F. nucleatum* loads in cancer patients' tissue corresponded to longer survival times; conversely, moderate and high bacterial loads corresponded to shorter survival times [38]. Another study also confirmed similar results and implied *F. nucleatum* might serve as a CRC prognosis marker [40].

As previous discussed, FadA and Fap2 are two virulence factors participating in tumorigenesis. FadA adhesion is highly conserved in *F. nucleatum*, and binds to E-cadherin on intestinal epithelial cells, inducing beta-cadherin signaling leading to Wnt, oncogene, and transcription factor expression, further promoting uncontrolled cell proliferation and inflammatory responses [21]. *In vivo* studies revealed that nude mice inoculated with CRC cells displayed a tumor mass increase of 20 % when followed with FadA treatment for 21 days post inoculation [21]. Additionally, FadA expression levels in CRC patient colon tissues is significantly higher than in healthy control [21]. Fap2 has been shown to actively bind Gal-Gal-NAc, a polysaccharide over expressed in human colorectal adenocarcinoma. This binding activity contributes to the abundance of *Fusobacteria* in CRC [22]. Fap2 directly actives TIGIT receptors, inhibitors of NK cell and T cell activity [21, 23]. Briefly, Fap2 is recognized by TIGIT receptor; Fap2/TIGIT binding inhibits NK cell cytotoxicity and T cell activity, avoiding immune attack. $Fap2^{mut}$ *F. nucleatum* is not able to bind to TIGIT, as indicated using hemagglutinin binding assays [23].

1.1.5 *F. nucleatum* in Other Human Diseases

In addition to periodontal disease and CRC, *F. nucleatum* also widely associate with other systemic diseases. Adverse pregnancy outcomes (APO), including preterm birth, preterm premature rupture of membranes, stillbirth, neonatal sepsis etc. was found to be correlated to periodontitis, and *F. nucleatum* was one of the predominant oral species found among placental and fetal tissues that were found to be closely associated with APO [41]. Animal studies indicated *F. nucleatum* migrated through hematogenous transmission from the oral cavity to the placenta, inducing APO [42]. Studies have revealed that the fetal loss rate was significantly reduced in *F. nucleatum* challenged Toll-like receptor 4 (TLR4) knockout mice compared with wild type mice, and accompanying necroinflammatory responses was also decreased, which suggests that *F. nucleatum* induced fetal death via a TLR4 modulated immune response [43-44]. *F. nucleatum* is also a major cause of Lemierre's Syndrome, a complication of pharyngitis in adolescents [45-46]. Additional diseases including inflammatory bowel disease, appendicitis, atherosclerosis, cerebral aneurysm, rheumatoid arthritis and Alzheimer's disease were also observed with *F. nucleatum* [15].

1.2 Colorectal Carcinoma

1.2.1 Epidemiology

The abnormal cell growth in the colon or rectum is regarded as CRC. It is the fourth leading cancer-caused death and third most common malignancy worldwide [47]. According to the statistic from World Health Organization in 2012, the incidence rate was highest in Korea following by Slovakia and some developed countries in Europe, whereas lowest rates occurred in Asian and African countries [48]. Epidemiological analysis indicated CRC rates was positively associated with the

human development index (HDI) [47]. The mortality rate was highest in Hungary, while lowest rates occurred in some countries in Africa [48]. While the trend of incidence and mortality varied, morbidity and mortality rates decreased in some high-income countries, probably due to better health care systems in place enabling early diagnosis and treatment [47, 49]. It was speculated that the CRC incidence and mortality rates are linked with HDI where [47, 49].

CRC typically presents no symptoms in the early stages, while as the tumor enlarges, obstruction occurs in the intestine. Risk factors associated with CRC are complex, and some factors like age, sex, family history, and inflammatory bowel disease are positively correlated with CRC risk. However, some medical factors including large bowel endoscopy, hormone replacement therapy, and aspirin are negatively related with CRC risk [49]. Rapidly evolving experimental techniques and accumulating lines of evidence show that *F. nucleatum, E.coli, B. fragilis* and other potential infectious bacteria may participate in CRC development and progression [50].

1.2.2 CRC Carcinogenesis and Metastasis

CRC typically starts from adenoma, a noncancerous growth, and often remains dormant in colon or rectum for decades [51]. Only fewer than 10% of the adenoma progress into cancer, though all adenomatous polyps have carcinogenic potential [52]. Once precursor adenomas leave the dormant stage they grow larger and develop into adenocarcinomas, infiltrating into the intestinal epithelium [49]. These masses are often metastatic, shedding cells that circulate to lymph nodes and other organs via blood vessels [49]. The molecular factors of CRC are heterogeneous though mutated oncogenes and tumor suppressor genes are assumed to play a role in adenocarcinogenesis [53]. In this process, KRAS oncogene is activated and tumor suppressor gene TP53 expression is suppressed [49, 53]. This

is accompanied by chromosomal instability, and the promotion of carcinogenesis [49, 53]. DNA repair mismatch is another mechanism leading to microsatellite instability, which accelerates CRC formation [49]. Additionally, inherited traits may play a role in the molecular pathogenesis in CRC [49]. CRC patients often develop liver, lung and brain metastases, though the mechanisms governing metastasis are not well characterized [54-56].

1.2.3 CRC Prevention and Therapy

Three levels of preventative measures are used to reduce probability of CRC development. Firstly, a healthy lifestyle and diet are seen as an essential for primary prevention. Avoiding some potential risk factors like smoking, obesity, and high consumption of red meat may greatly reduce CRC incidence[49]. Studies also suggested chemoprevention supplements such as aspirin and vitamin D can reduce CRC [49]. A secondary preventive measure is early diagnosis and screening, which includes visual examination and stool tests. Colonoscopy is one typical and recommend method for visual examination that can reduce incidence and mortality rates significantly [58]. The fecal occult blood test is a highly specific but insensitive stool screen, which is a non-invasive and valuable measure for reducing CRC [59-60]. A third preventive category is risk reducing lifestyle choices. Exercise and smoking cessation are proven behaviors, which enhance general health and may significantly reduce CRC incidence[49]. Surgery is a standard and main treatment for non-metastasized CRC. Adjuvant therapy is recommended for stage III CRC and some other high-risk cases [49, 61].

1.3 Bacterial Outer Membrane Vesicles

1.3.1 Outer Membrane Vesicles Biogenesis

Outer membrane vesicles (OMVs) are a powerful new drug and vaccine delivery system. OMVs are spontaneously produced by Gram-negative bacterium as a result of normal bacterial metabolism [62]. Gram-negative bacteria possess two membranes, the outer and inner (or cytoplasmic) membranes; in between lies a periplasmic fraction mainly contain peptidoglycan [63]. OMVs disassociate from the bacterial outer membrane and are secreted outward. OMVs are loaded with outer membrane proteins and consist of phospholipid and lipopolysaccharide (LPS), and some periplasmic proteins. Although the OMVs were originally identified in lysed bacterial preparations, later studies demonstrated these vesicles to be present without cell disruption [62]. It has been proven the biogenesis of OMVs cargo is selected and premeditated, in response to environmental stresses including nutrient shortfall, oxidation, antibiotics, and others [64]. Additionally, OMVs isolated from a given batch or culture may contain a wide variety of different components [62]. OMV biogenesis starts when outer membrane bulges certain localized section form, and detach from the underlying peptidoglycan layer [65]. After bulge formation, vesicles are released spontaneously from the outer membrane layer [65]. The vesiculation process is very random; components of the peptidoglycan layer, portions of the inner membrane, and even contents of the cytosol located near a budding vesicle may become incorporated as vesicle contents or cargo.

1.3.2 Composition and Functions

The vesiculation process is a nearly universal process seen in Gram-negative bacteria; even in mutant strains, OMVs production also occurs [65]. Analyzing these vesicles is challenging, however,

sophisticated analytical tools have been recently developed such as proteomics. Several studies have shown that the bulk of OMVs consists of outer membrane, periplasmic, cytoplasmic and cytoplasmic membrane proteins [62, 66-67]. A few OMVs are bi-layered structures, with an inner membrane as well as an outer one; presumably, these layers correspond with the bacterial inner and outer membranes, respectively [68]. Multiple OMV biogenesis models have been proposed, but details concerning OMV biogenesis remain unclear; OMVs containing outer membranes only predominate, yet, in a small percentage of OMVs sections of the bacterial inner membrane may become incorporated into the budding process that leads to the generation of OMVs.

OMVs are known to play roles in bacterial pathogenesis; variousvirulence factors are often incorporated into OMVs and transported by these. VirD4, secreted via a type IV secretion system, is often found in OMVs of *Helicobacter pylori*; blood group antigen-binding adhesin (BabA) and sialic acid-binding adhesin (SabA) are molecules that facilitate the interaction with host epithelium and are also carried in OMVs [62, 69]. HtrA, a protease involving in misfolded-protein accumulation identified in OMVs secreted from both *E. coli* and *H. pylori*, and plays a roles in pathogenesis by disrupting E-cadherin cleavage [69-70]. Environmental stressors, such as temperature change and limited nutrients supply can trigger OMV upregulation in conjunction with stress related proteins, which become incorporated with the OMVs, altering OMV protein profiles [62]. OMVs serve a protective role in bacterial survival, mediating biofilm formation and antibiotic resistance [71]. OMVs are also a potent contributor of bacterial aggregation and enhance bacterium-bacterium interactions regarding quorum sensing and communication [72-73].

1.3.3 Role in Immune System

The immunogenicity prosperities of OMVs have been studied extensively in many bacterial strains; OMVs have been shown to interact with diverse host cells to modulate immune responses and protect bacteria from host immune attack. OMVs are secreted by both pathogenic and commensal bacteria during growth periods in a non-replicative vesicle form. Epithelial cells are the first defense line in the host, and are where OMVs typically translocate initially following secretion. OMVs collected from *H. pylori* can interact with gastric epithelium cells to promote dose-dependent interleukin-8 (IL-8) response [74]. Likewise, *Pseudomonas aeruginosa* OMVs engage in lung epithelial cells and stimulate IL-8 secretion [75]. *Klebsiella pneumoniae* OMVs stimulate IL-1β and IL-8 production in laryngeal epithelial cells [66, 76]. An *in vivo* study has reported that *H. pylori* OMVs containing peptidoglycans that can be recognized by nucleotide-binding oligomerization domain-containing protein 1 (NOD1) in the host. The binding of which presumably induces innate and adaptive immune responses [77].

TLR4 is a well-studied pattern recognition receptor (PRR) that often interacts with OMVs. LPS presenting on OMVs from *E. coli* arouses TLR4 dependent immune responses [78]. Another study shows that OMVs mediating tumor necrosis factor (TNF) tolerance in monocytes through TLR4 [79]. In addition to colonizing surface epithelial cells, OMVs can induce physiological damage to epithelium, allowing OMV transported virulence factors to interact with cells in the submucosa[80]. After transit into underlying tissues, OMVs interact with multiple immune cells to induce pro-inflammatory activity. *N. meningitides* OMVs activate neutrophils, triggering cytokines including TNFα and IL-1β; *E. coli* OMVs contain cytotoxic necrotizing factor type 1 that restrain antimicrobial activity and neutrophil chemotaxis [81-82]. OMVs interact with

macrophages cytokines, for example, *N. meningitidis* OMVs induce monocyte and macrophage production of IL-1β, IL-6 and other cytokines [83] (Table 1). Meanwhile, these interactions with macrophages also trigger adaptive immune responses, stimulating upregulation of HLA-DR, CD80, and CD86 [83]. OMVs are able to promote dendritic cells (DCs) maturation and foster B cell and T cell responses [84]. When OMVs circulate to other body sites in addition to the primary infection area, they can modulate other host cells including platelets and endothelial cells [85-86]. Overall, these immunogenic abilities make OMVs ideal components for vaccines and adjuvants.

Table 1 Role of OMVs in immune system

Bacteria	**Function in immune system**
H. pylori	Promote IL-8 response; target NOD1 and induce innate response
P. aeruginosa	Encourage IL-8 secretion
E. coli	Present LPS and stimulate TLR4 response; Restrain antimicrobial activity and chemotaxis of netrophil
K. pneumoniae	Stimulate IL-1β and IL-8
N. meningitides	Stimulate neutrophil to produce cytokines; Inspire monocyte and macrophage to produce IL-1β, IL-6 and other cytokines

1.3.4 Vaccine Candidates and Application

Stable, non-replicated nano-sized properties of OMVs give them great potential for vaccine design. The proteomic study by mass spectrometry and computational analysis classified components of OMVs from different strains; multiple molecules have been postulated as vaccine candidates.

An OMV preparation developed against *N. meningitides* serogroup B is the first approved vaccine generated from the wild type strain [87-88]. ProA and ProB are the two main prion antigens transported by OMVs, while other minor proteins have been used for cross-protection [89-90]. For example, Bexsero® using OMVs with capsular polysaccharide combined with a recombinant protein have been developed and proven as a safe vaccine [91]. OMVs from a genetically engineered strain with attenuated LPS toxicity and overexpressed PorA provide a safer and more effective vaccine approach [92]. OMVs isolated from epidemic strains against serogroups A, W and X, which cause meningitis in Africa, have been examined under clinical trials for vaccine development [88].

Bordetella pertussis is the major cause of whooping cough in both children and adults. Roberts *et al*. identified 43 proteins in the OMVs of *B. pertussis*. Among these, numerous major surface antigens such as adenylate cyclase-hemolysin, pertactin, and lipo-oligosaccharide were found [93]. Significant immune and protective responses were observed in mice immunized with the above [93]. Engineered OMVs collected from inactive *B. pertussis* carried multiple antigens including pertussis toxin and a lipid A 3-deacylase have been evaluated in an intranasal challenge model and were shown to induce protection [94]. This study supports the speculation that OMVs from *B. pertussis* have strong potential as a candidate for vaccine against pertussis.

Other investigations of OMV-based vaccine approach against human diseases, including enteric disease and tuberculosis, have been developed in the past decade. OMVs extracted from *V. cholera, Salmonella* and *Shigella*, the main pathogens for enteric infection, are able to trigger extensive IgG and IgA level responses [88]. LPS-modified OMVs collected from *V. cholera* and *E. coli* have the ability to elicit specific protective immune responses and have been explored as a heterologous vaccine delivery platform [95]. OMVs from *Mycobacteria* have been verified as interacting with TLR2 to stimulate cytokine and chemokine response in a mouse model, which provides evidence as a vaccine candidate for eliciting active immune responses [96].

Overall, OMVs are expected to become widely used as an advanced platform for safer and well-mediated vaccines against diverse pathogenic and commensal bacterial strains associated with various human diseases.

1.4 Vaccine Development

The first vaccine was developed in Asia 1,000 years ago against the smallpox virus. Later in the eighteenth century, Edward Jenner noticed milkmen infected with cowpox acquired the ability against smallpox infection. since then, vaccinology has changed and developed with the aim of generating vaccines that are both highly immunogenic and present few side effects. Two major methods have been established to enhance vaccine immunogenicity: attenuated pathogens and recombinant proteins. There are numerous dilemmas in both approaches. In the first one, immunogenicity may compromise for safety and tolerability; and the second approach is limited by side effects. Therefore, an integrated approach to design vaccine capable of mimicking properties of pathogens without causing disease is necessary.

Important factors include the size, geometry, kinetics and surface antigens[97]. The size of vaccine determines the available of uptake by antigen presenting cells (APCs), which is crucial for T cell response. It has been speculated antigens around 20-100nm in size (similar to pathogen size), are able to maximize the uptake capacity of APCs or transport into lymph vessels [97]. The geometry and the quantity of antigenic determinants of the vaccine is essential for B and T cell responses. Highly repetitive particle patterns similar to the spacing of pathogen coat proteins are crucial for triggering B cell responses [97-98]. The surface pattern is significant in optimizing long-lived T helper cell responses. Additionally, subsequent boosters instead of single dose injection might prolong and enhance the protective efficacy.

In summary, nanoparticle vaccines, such as OMVs, have proven themselves to be very useful in vaccine design. OMVs as a type of nanoparticle with the proper size between 20-200 nm and are not replicable, making them safe as a vaccine delivery platform. Furthermore, OMVs are isolated from cultures of pathogens, and therefore antigenically similar if not identical to the pathogen under consideration. These qualities allow OMV vaccine preparations to mimic pathogen specifically and precisely, thus are able to elicit specific immune responses. OMVs will undoubtedly become the basis for a whole generation of upcoming vaccine designs.

REFERENCES

1. Sender R, Fuchs S, Milo R. Revised Estimates for the Number of Human and Bacteria Cells in the Body. PLoS Biol. 2016;14(8):e1002533.
2. Klimesova K, Jiraskova Zakostelska Z, Tlaskalova-Hogenova H. Oral Bacterial and Fungal Microbiome Impacts Colorectal Carcinogenesis. Front Microbiol. 2018;9:774.
3. Charbonneau MR, Blanton LV, DiGiulio DB, Relman DA, Lebrilla CB, Mills DA, et al. A microbial perspective of human developmental biology. Nature. 2016;535(7610):48-55.
4. Schroeder BO, Backhed F. Signals from the gut microbiota to distant organs in physiology and disease. Nat Med. 2016;22(10):1079-89.
5. Schaffler H, Breitruck A. *Clostridium difficile* - From Colonization to Infection. Front Microbiol. 2018;9:646.
6. Niederreiter L, Adolph TE, Tilg H. Food, microbiome and colorectal cancer. Dig Liver Dis. 2018.
7. Guinane CM, Cotter PD. Role of the gut microbiota in health and chronic gastrointestinal disease: understanding a hidden metabolic organ. Therap Adv Gastroenterol. 2013;6(4):295-308.
8. Bolstad AI, Jensen HB, Bakken V. Taxonomy, biology, and periodontal aspects of *Fusobacterium nucleatum*. Clin Microbiol Rev. 1996;9(1):55-71.
9. Isomura K. Bacterial endotoxin from *Fusobacterium nucleatum*--particularly the role of endotoxin in periodontal diseases. Nihon Shishubyo Gakkai Kaishi. 1985;27(2):328-51.
10. Strauss J, White A, Ambrose C, McDonald J, Allen-Vercoe E. Phenotypic and genotypic analyses of clinical *Fusobacterium nucleatum* and *Fusobacterium periodonticum* isolates from the human gut. Anaerobe. 2008;14(6):301-9.
11. Dzink JL, Sheenan MT, Socransky SS. Proposal of three subspecies of *Fusobacterium nucleatum* Knorr 1922: *Fusobacterium nucleatum subsp. nucleatum subsp. nov., comb. nov.*; *Fusobacterium nucleatum subsp. polymorphum subsp. nov., nom. rev., comb. nov.*; and *Fusobacterium nucleatum subsp. vincentii subsp. nov., nom. rev., comb. nov.* Int J Syst Bacteriol. 1990;40(1):74-8.
12. Spaulding EH, Rettger LF. The *Fusobacterium* Genus: I. Biochemical and Serological Classification. J Bacteriol. 1937;34(5):535-48.
13. Gharbia SE, Shah HN. *Fusobacterium nucleatum subsp. fusiforme subsp. nov.* and *Fusobacterium nucleatum subsp. animalis subsp. nov.* as additional subspecies within *Fusobacterium nucleatum*. Int J Syst Bacteriol. 1992;42(2):296-8.
14. Citron DM. Update on the taxonomy and clinical aspects of the genus *fusobacterium*. Clin Infect Dis. 2002;35(Suppl 1):S22-7.
15. Han YW. *Fusobacterium nucleatum*: a commensal-turned pathogen. Curr Opin Microbiol. 2015;23:141-7.
16. Williams MD, Kerber CA, Tergin HF. Unusual presentation of Lemierre's syndrome due to *Fusobacterium nucleatum*. J Clin Microbiol. 2003;41(7):3445-8.
17. Sparks Stein P, Steffen MJ, Smith C, Jicha G, Ebersole JL, Abner E, et al. Serum antibodies to periodontal pathogens are a risk factor for Alzheimer's disease. Alzheimers Dement. 2012;8(3):196-203.

18. Kaplan CW, Lux R, Haake SK, Shi W. The *Fusobacterium nucleatum* outer membrane protein RadD is an arginine-inhibitable adhesin required for inter-species adherence and the structured architecture of multispecies biofilm. Mol Microbiol. 2009;71(1):35-47.
19. Coppenhagen-Glazer S, Sol A, Abed J, Naor R, Zhang X, Han YW, et al. Fap2 of *Fusobacterium nucleatum* is a galactose-inhibitable adhesin involved in coaggregation, cell adhesion, and preterm birth. Infect Immun. 2015;83(3):1104-13.
20. Xu M, Yamada M, Li M, Liu H, Chen SG, Han YW. FadA from *Fusobacterium nucleatum* utilizes both secreted and nonsecreted forms for functional oligomerization for attachment and invasion of host cells. J Biol Chem. 2007;282(34):25000-9.
21. Rubinstein MR, Wang X, Liu W, Hao Y, Cai G, Han YW. *Fusobacterium nucleatum* promotes colorectal carcinogenesis by modulating E-cadherin/beta-catenin signaling via its FadA adhesin. Cell Host Microbe. 2013;14(2):195-206.
22. Abed J, Emgard JE, Zamir G, Faroja M, Almogy G, Grenov A, et al. Fap2 Mediates *Fusobacterium nucleatum* Colorectal Adenocarcinoma Enrichment by Binding to Tumor-Expressed Gal-GalNAc. Cell Host Microbe. 2016;20(2):215-25.
23. Gur C, Ibrahim Y, Isaacson B, Yamin R, Abed J, Gamliel M, et al. Binding of the Fap2 protein of *Fusobacterium nucleatum* to human inhibitory receptor TIGIT protects tumors from immune cell attack. Immunity. 2015;42(2):344-55.
24. Kuboniwa M, Lamont RJ. Subgingival biofilm formation. Periodontol 2000. 2010;52(1):38-52.
25. Yang NY, Zhang Q, Li JL, Yang SH, Shi Q. Progression of periodontal inflammation in adolescents is associated with increased number of *Porphyromonas gingivalis*, *Prevotella intermedia, Tannerella forsythensis*, and *Fusobacterium nucleatum*. Int J Paediatr Dent. 2014;24(3):226-33.
26. Field CA, Gidley MD, Preshaw PM, Jakubovics N. Investigation and quantification of key periodontal pathogens in patients with type 2 diabetes. J Periodontal Res. 2012;47(4):470-8.
27. Griffen AL, Beall CJ, Campbell JH, Firestone ND, Kumar PS, Yang ZK, et al. Distinct and complex bacterial profiles in human periodontitis and health revealed by 16S pyrosequencing. Isme J. 2012;6(6):1176-85.
28. van Palenstein Helderman WH. Total viable count and differential count of *Vibrio (campylobacter) sputorum, Fusobacterium nucleatum, Selenomonas sputigena, Bacteroides ochraceus* and *Veillonella* in the inflamed and non inflamed human gingival crevice. J Periodontal Res. 1975;10(5):294-305.
29. Falkler WA, Jr., Lai R, Vincent JW, Dober L, Spiegel C, Hayduk S. The ELISA system for measuring antibody reactive to *Fusobacterium nucleatum* in the sera of patients with chronic periodontitis. J Periodontol. 1982;53(12):762-6.
30. Baumgartner JC, Falkler WA, Jr., Beckerman T. Experimentally induced infection by oral anaerobic microorganisms in a mouse model. Oral Microbiol Immunol. 1992;7(4):253-6.
31. Kinder SA, Holt SC. Characterization of coaggregation between *Bacteroides gingivalis* T22 and *Fusobacterium nucleatum* T18. Infect Immun. 1989;57(11):3425-33.
32. Kolenbrander PE, Andersen RN. Inhibition of coaggregation between *Fusobacterium nucleatum* and *Porphyromonas (Bacteroides) gingivalis* by lactose and related sugars. Infect Immun. 1989;57(10):3204-9.

33. Kaufman J, DiRienzo JM. Evidence for the existence of two classes of corncob (coaggregation) receptor in *Fusobacterium nucleatum*. Oral Microbiol Immunol. 1988;3(4):145-52.
34. Lancy P, Jr., Dirienzo JM, Appelbaum B, Rosan B, Holt SC. Corncob formation between *Fusobacterium nucleatum* and *Streptococcus sanguis*. Infect Immun. 1983;40(1):303-9.
35. Kaufman J, DiRienzo JM. Isolation of a corncob (coaggregation) receptor polypeptide from *Fusobacterium nucleatum*. Infect Immun. 1989;57(2):331-7.
36. Diaz PI, Zilm PS, Rogers AH. *Fusobacterium nucleatum* supports the growth of *Porphyromonas gingivalis* in oxygenated and carbon-dioxide-depleted environments. Microbiol-Sgm. 2002;148:467-72.
37. Castellarin M, Warren RL, Freeman JD, Dreolini L, Krzywinski M, Strauss J, et al. *Fusobacterium nucleatum* infection is prevalent in human colorectal carcinoma. Genome Res. 2012;22(2):299-306.
38. Flanagan L, Schmid J, Ebert M, Soucek P, Kunicka T, Liska V, et al. *Fusobacterium nucleatum* associates with stages of colorectal neoplasia development, colorectal cancer and disease outcome. Eur J Clin Microbiol Infect Dis. 2014;33(8):1381-90.
39. Kostic AD, Chun E, Robertson L, Glickman JN, Gallini CA, Michaud M, et al. *Fusobacterium nucleatum* potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. Cell Host Microbe. 2013;14(2):207-15.
40. Mima K, Nishihara R, Qian ZR, Cao Y, Sukawa Y, Nowak JA, et al. *Fusobacterium nucleatum* in colorectal carcinoma tissue and patient prognosis. Gut. 2016;65(12):1973-80.
41. Gonzales-Marin C, Spratt DA, Allaker RP. Maternal oral origin of *Fusobacterium nucleatum* in adverse pregnancy outcomes as determined using the 16S-23S rRNA gene intergenic transcribed spacer region. J Med Microbiol. 2013;62(Pt 1):133-44.
42. Stockham S, Stamford JE, Roberts CT, Fitzsimmons TR, Marchant C, Bartold PM, et al. Abnormal pregnancy outcomes in mice using an induced periodontitis model and the haematogenous migration of Fusobacterium nucleatum sub-species to the murine placenta. PLoS One. 2015;10(3):e0120050.
43. Liu H, Redline RW, Han YW. *Fusobacterium nucleatum* induces fetal death in mice via stimulation of TLR4-mediated placental inflammatory response. J Immunol. 2007;179(4):2501-8.
44. Vander Haar EL, So J, Gyamfi-Bannerman C, Han YW. *Fusobacterium nucleatum* and adverse pregnancy outcomes: Epidemiological and mechanistic evidence. Anaerobe. 2018;50:55-9.
45. Brook I. Fusobacterial head and neck infections in children. Int J Pediatr Otorhinolaryngol. 2015;79(7):953-8.
46. Brook I. *Fusobacterial* infections in children. Curr Infect Dis Rep. 2013;15(3):288-94.
47. Arnold M, Sierra MS, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global patterns and trends in colorectal cancer incidence and mortality. Gut. 2017;66(4):683-91.
48. Organization WH. GLOBOCAN 2012 [Available from: http://gco.iarc.fr/today/online-analysis-map?mode=population&mode_population=continents&population=900&sex=0&cancer=6&type=0&statistic=2&prevalence=0&color_palette=default&projection=natural-earth.
49. Brenner H, Kloor M, Pox CP. Colorectal cancer. Lancet. 2014;383(9927):1490-502.

50. Tilg H, Adolph TE, Gerner RR, Moschen AR. The Intestinal Microbiota in Colorectal Cancer. Cancer Cell. 2018;33(6):954-64.
51. Siegel RL, Miller KD, Fedewa SA, Ahnen DJ, Meester RGS, Barzi A, et al. Colorectal cancer statistics, 2017. CA Cancer J Clin. 2017;67(3):177-93.
52. Levine JS, Ahnen DJ. Clinical practice. Adenomatous polyps of the colon. N Engl J Med. 2006;355(24):2551-7.
53. Ewing I, Hurley JJ, Josephides E, Millar A. The molecular genetics of colorectal cancer. Frontline Gastroenterol. 2014;5(1):26-30.
54. Zang YW, Gu XD, Xiang JB, Chen ZY. Brain metastases from colorectal cancer: microenvironment and molecular mechanisms. Int J Mol Sci. 2012;13(12):15784-800.
55. Akgul O, Cetinkaya E, Ersoz S, Tez M. Role of surgery in colorectal cancer liver metastases. World J Gastroenterol. 2014;20(20):6113-22.
56. Kim HS, Heo JS, Lee J, Lee JY, Lee MY, Lim SH, et al. The impact of KRAS mutations on prognosis in surgically resected colorectal cancer patients with liver and lung metastases: a retrospective analysis. BMC Cancer. 2016;16:120.
57. Singh Ranger G. The role of aspirin in colorectal cancer chemoprevention. Crit Rev Oncol Hematol. 2016;104:87-90.
58. Brenner H, Chang-Claude J, Seiler CM, Rickert A, Hoffmeister M. Protection from colorectal cancer after colonoscopy: a population-based, case-control study. Ann Intern Med. 2011;154(1):22-30.
59. Kawabata H, Inoue N, Kawakatsu Y, Okazaki Y, Hitomi M, Miyata M, et al. Screening for colorectal cancer using immunological fecal occult blood test in inpatients. Nihon Shokakibyo Gakkai Zasshi. 2018;115(4):377-84.
60. Zubero MB, Arana-Arri E, Pijoan JI, Portillo I, Idigoras I, Lopez-Urrutia A, et al. Population-based colorectal cancer screening: comparison of two fecal occult blood test. Front Pharmacol. 2014;4:175.
61. Kuipers EJ, Grady WM, Lieberman D, Seufferlein T, Sung JJ, Boelens PG, et al. Colorectal cancer. Nat Rev Dis Primers. 2015;1:15065.
62. Schwechheimer C, Kuehn MJ. Outer-membrane vesicles from Gram-negative bacteria: biogenesis and functions. Nat Rev Microbiol. 2015;13(10):605-19.
63. Silhavy TJ, Kahne D, Walker S. The bacterial cell envelope. Cold Spring Harb Perspect Biol. 2010;2(5):a000414.
64. Macdonald IA, Kuehn MJ. Stress-induced outer membrane vesicle production by *Pseudomonas aeruginosa*. J Bacteriol. 2013;195(13):2971-81.
65. Kulp A, Kuehn MJ. Biological functions and biogenesis of secreted bacterial outer membrane vesicles. Annu Rev Microbiol. 2010;64:163-84.
66. Aguilera L, Toloza L, Gimenez R, Odena A, Oliveira E, Aguilar J, et al. Proteomic analysis of outer membrane vesicles from the probiotic strain *Escherichia coli Nissle 1917*. Proteomics. 2014;14(2-3):222-9.
67. Pierson T, Matrakas D, Taylor YU, Manyam G, Morozov VN, Zhou W, et al. Proteomic characterization and functional analysis of outer membrane vesicles of *Francisella novicida* suggests possible role in virulence and use as a vaccine. J Proteome Res. 2011;10(3):954-67.
68. Perez-Cruz C, Delgado L, Lopez-Iglesias C, Mercade E. Outer-inner membrane vesicles naturally secreted by gram-negative pathogenic bacteria. PLoS One. 2015;10(1):e0116896.

69. Olofsson A, Vallstrom A, Petzold K, Tegtmeyer N, Schleucher J, Carlsson S, et al. Biochemical and functional characterization of *Helicobacter pylori* vesicles. Mol Microbiol. 2010;77(6):1539-55.
70. Hoy B, Geppert T, Boehm M, Reisen F, Plattner P, Gadermaier G, et al. Distinct roles of secreted HtrA proteases from gram-negative pathogens in cleaving the junctional protein and tumor suppressor E-cadherin. J Biol Chem. 2012;287(13):10115-20.
71. Murphy K, Park AJ, Hao Y, Brewer D, Lam JS, Khursigara CM. Influence of O polysaccharides on biofilm development and outer membrane vesicle biogenesis in *Pseudomonas aeruginosa* PAO1. J Bacteriol. 2014;196(7):1306-17.
72. Kulkarni HM, Jagannadham MV. Biogenesis and multifaceted roles of outer membrane vesicles from Gram-negative bacteria. Microbiology. 2014;160(Pt 10):2109-21.
73. Mashburn-Warren L, Howe J, Garidel P, Richter W, Steiniger F, Roessle M, et al. Interaction of quorum signals with outer membrane lipids: insights into prokaryotic membrane vesicle formation. Mol Microbiol. 2008;69(2):491-502.
74. Ismail S, Hampton MB, Keenan JI. Helicobacter pylori outer membrane vesicles modulate proliferation and interleukin-8 production by gastric epithelial cells. Infect Immun. 2003;71(10):5670-5.
75. Bauman SJ, Kuehn MJ. Purification of outer membrane vesicles from *Pseudomonas aeruginosa* and their activation of an IL-8 response. Microbes Infect. 2006;8(9-10):2400-8.
76. Lee JC, Lee EJ, Lee JH, Jun SH, Choi CW, Kim SI, et al. *Klebsiella pneumoniae* secretes outer membrane vesicles that induce the innate immune response. FEMS Microbiol Lett. 2012;331(1):17-24.
77. Kaparakis M, Turnbull L, Carneiro L, Firth S, Coleman HA, Parkington HC, et al. Bacterial membrane vesicles deliver peptidoglycan to NOD1 in epithelial cells. Cell Microbiol. 2010;12(3):372-85.
78. Soderblom T, Oxhamre C, Wai SN, Uhlen P, Aperia A, Uhlin BE, et al. Effects of the *Escherichia coli* toxin cytolysin A on mucosal immunostimulation via epithelial Ca2+ signalling and Toll-like receptor 4. Cell Microbiol. 2005;7(6):779-88.
79. Waller T, Kesper L, Hirschfeld J, Dommisch H, Kolpin J, Oldenburg J, et al. *Porphyromonas gingivalis* Outer Membrane Vesicles Induce Selective Tumor Necrosis Factor Tolerance in a Toll-Like Receptor 4- and mTOR-Dependent Manner. Infect Immun. 2016;84(4):1194-204.
80. Nakao R, Takashiba S, Kosono S, Yoshida M, Watanabe H, Ohnishi M, et al. Effect of *Porphyromonas gingivalis* outer membrane vesicles on gingipain-mediated detachment of cultured oral epithelial cells and immune responses. Microbes Infect. 2014;16(1):6-16.
81. Davis JM, Carvalho HM, Rasmussen SB, O'Brien AD. Cytotoxic necrotizing factor type 1 delivered by outer membrane vesicles of uropathogenic *Escherichia coli* attenuates polymorphonuclear leukocyte antimicrobial activity and chemotaxis. Infect Immun. 2006;74(8):4401-8.
82. Lapinet JA, Scapini P, Calzetti F, Perez O, Cassatella MA. Gene expression and production of tumor necrosis factor alpha, interleukin-1beta (IL-1beta), IL-8, macrophage inflammatory protein 1alpha (MIP-1alpha), MIP-1beta, and gamma interferon-inducible protein 10 by human neutrophils stimulated with group B *meningococcal* outer membrane vesicles. Infect Immun. 2000;68(12):6917-23.

83. Tavano R, Franzoso S, Cecchini P, Cartocci E, Oriente F, Arico B, et al. The membrane expression of *Neisseria meningitidis* adhesin A (NadA) increases the proimmune effects of MenB OMVs on human macrophages, compared with NadA- OMVs, without further stimulating their proinflammatory activity on circulating monocytes. J Leukoc Biol. 2009;86(1):143-53.
84. Durand V, Mackenzie J, de Leon J, Mesa C, Quesniaux V, Montoya M, et al. Role of lipopolysaccharide in the induction of type I interferon-dependent cross-priming and IL-10 production in mice by meningococcal outer membrane vesicles. Vaccine. 2009;27(13):1912-22.
85. Mirlashari MR, Hagberg IA, Lyberg T. Platelet-platelet and platelet-leukocyte interactions induced by outer membrane vesicles from *N. meningitidis*. Platelets. 2002;13(2):91-9.
86. Kim JH, Yoon YJ, Lee J, Choi EJ, Yi N, Park KS, et al. Outer membrane vesicles derived from *Escherichia coli* up-regulate expression of endothelial cell adhesion molecules in vitro and in vivo. PLoS One. 2013;8(3):e59276.
87. Holst J, Oster P, Arnold R, Tatley MV, Naess LM, Aaberge IS, et al. Vaccines against *meningococcal* serogroup B disease containing outer membrane vesicles (OMV): lessons from past programs and implications for the future. Hum Vaccin Immunother. 2013;9(6):1241-53.
88. Acevedo R, Fernandez S, Zayas C, Acosta A, Sarmiento ME, Ferro VA, et al. Bacterial outer membrane vesicles and vaccine applications. Front Immunol. 2014;5:121.
89. Feavers IM, Pizza M. Meningococcal protein antigens and vaccines. Vaccine. 2009;27 Suppl 2:B42-50.
90. Holst J, Martin D, Arnold R, Huergo CC, Oster P, O'Hallahan J, et al. Properties and clinical performance of vaccines containing outer membrane vesicles from N*eisseria meningitidis*. Vaccine. 2009;27 Suppl 2:B3-12.
91. Bai X, Findlow J, Borrow R. Recombinant protein *meningococcal* serogroup B vaccine combined with outer membrane vesicles. Expert Opin Biol Ther. 2011;11(7):969-85.
92. van de Waterbeemd B, Streefland M, van der Ley P, Zomer B, van Dijken H, Martens D, et al. Improved OMV vaccine against *Neisseria meningitidis* using genetically engineered strains and a detergent-free purification process. Vaccine. 2010;28(30):4810-6.
93. Roberts R, Moreno G, Bottero D, Gaillard ME, Fingermann M, Graieb A, et al. Outer membrane vesicles as acellular vaccine against pertussis. Vaccine. 2008;26(36):4639-46.
94. Asensio CJ, Gaillard ME, Moreno G, Bottero D, Zurita E, Rumbo M, et al. Outer membrane vesicles obtained from *Bordetella pertussis Tohama* expressing the lipid A deacylase PagL as a novel acellular vaccine candidate. Vaccine. 2011;29(8):1649-56.
95. Leitner DR, Lichtenegger S, Temel P, Zingl FG, Ratzberger D, Roier S, et al. A combined vaccine approach against *Vibrio cholerae* and ETEC based on outer membrane vesicles. Front Microbiol. 2015;6:823.
96. Prados-Rosales R, Baena A, Martinez LR, Luque-Garcia J, Kalscheuer R, Veeraraghavan U, et al. *Mycobacteria* release active membrane vesicles that modulate immune responses in a TLR2-dependent manner in mice. J Clin Invest. 2011;121(4):1471-83.
97. Bachmann MF, Jennings GT. Vaccine delivery: a matter of size, geometry, kinetics and molecular patterns. Nat Rev Immunol. 2010;10(11):787-96.
98. Bachmann MF, Zinkernagel RM. Neutralizing antiviral B cell responses. Annu Rev Immunol. 1997;15:235-70.

CHAPTER TWO

PROTEOMIC CHARACTERIZATION OF OUTER MEMBRANE VESICLES FROM GUT MUCOSA-DERIVED *FUSOBACTERIUM NUCLEATUM*

Introduction

Fusobacterium nucleatum is an opportunistic Gram-negative non-spore forming and obligate anaerobic bacterium with a fusiform rod shape and is a key component of plaque and a common resident of human oral cavity [1-3]. It is a heterogeneous species and one of the most abundant organisms identified in endodontic infections [1, 3]. *F. nucleatum* is a primary contributor to many types of periodontal diseases [4-5]. It is abundant in the oral cavity of patients with periodontitis and gingivitis, although it can also be detected in healthy people at lower levels [5-8]. Chronic periodontal disease patients often show elevated IgG and IgA responses to *F. nucleatum*, whose bacterial load has been positively correlated with the progression of disease [5-6, 9]. *F. nucleatum* can reduce the oxygen concentration in the microenvironment, thus facilitating the growth of other anaerobic bacteria, such as *Porphyromonas gingivalis* [10]. These mixed anaerobic colonies form coaggregation units and play an essential role in cell surface attachment and plaque formation in the oral cavity [4, 11-13].

Recent lines of evidence have suggested an oncogenic role of *F. nucleatum* in the development of colorectal tumors [14]. The human gut contains a great variety of microbial species, some of which have long been suspected of being potentiators of colorectal cancer (CRC), the fourth most frequent cause of cancer death worldwide [15-16]. More recently, studies have linked gut colonization with the oral cavity-resident bacterium, *F. nucleatum*, with CRC tumorigenesis and metastasis [17].

This association led many researchers to postulate that *F. nucleatum* is a contributor to intestinal carcinogenesis [18-19].

In multiple studies, *F. nucleatum* abundance consistently increases comparing normal colon to adenoma and adenoma to CRC [18, 20-23]. Higher CRC *F. nucleatum* levels are associated with increased WNT and NFKB signaling, production of IL17 and other cytokines, tumor associated myeloid cells, fewer tumor infiltrating T cells (TILs) and shorter survival[24-25]. *In vitro*, *F. nucleatum* co-culture with conventional CRC cell lines increases tumor promoting autophagy, TLR4 and RIG-1 signaling [21-23]. The *F. nucleatum* adhesin FadA binds to CDH1 and promotes colonocyte WNT/proliferation [26]; and the lectin Fap2 binds to and promotes adhesion to human TIGIT, which inhibits recruitment of tumor-infiltrating NK cells and lymphocytes [27]. In *Apc* mutant mice, oral gavage with *F. nucleatum* increases small intestinal and colon adenomas (4). *F. nucleatum* levels in CRC are most strongly associated with right-sided location, mucinous adenocarcinomas and mismatch repair deficient CRC [18, 20, 24-25].

Many colorectal tumor tissue samples have been shown to have relatively heavy *F. nucleatum* bacterial loads compared with adjacent normal intestinal tissue sites [18, 28]. One well-characterized virulence factors of *F. nucleatum* is the adhesion protein FadA, which facilitates bacterial attachment and invasion [26, 29]. FadA promotes carcinogenesis via binding to an E-cadherin receptor, activating β-catenin and driving up expression of transcriptional factors, Wnt genes, inflammatory genes, and associated oncogenes [26]. Another associated virulence factor of *F. nucleatum* is lectin Fap2, a galactose adhesion protein, which binds Gal-GalNAc, facilitating localization and colonization of *F. nucleatum* in CRC cells [27]. Additionally, Fap2 suppresses T cell and natural

killer (NK) cell activities via inhibitory receptor TIGIT binding; this results in the impairment of T cell/NK cell activity, allowing cancerous tumors to effectively evade immune responses [21].

The most commonly practiced prevention strategy for CRC is chemoprevention such as using aspirin and other nonsteroidal anti-inflammatory drugs (NSAIDs) [30-31]. However, long-term use of aspirin (or other NSAIDs) is associated with adverse effects including gastrointestinal (GI) bleeding. Safer alternatives for CRC prevention are urgently needed. Another potential approach to reduce *F. nucleatum* colonization is vaccination. Outer membrane vesicles (OMVs) have recently emerged as powerful vaccine delivery systems. OMVs are naturally secreted, bi-layered lipid membrane nanostructures produced by Gram-negative bacteria as metabolic byproducts [32]. OMVs offer characteristics that are ideal for vaccine applications. For example, they contain much of the biological content of the parent bacterium, but without replicative or infectious capacity; are highly stable and highly immunogenic; and can interact with epithelial and immune cells in content dependent manner. In 2013, the European Medicines Agency (EMA) and US Food and Drug Administration (FDA) approved a vaccine containing OMV components for inoculation against infection by meningococcal serogroup B [33]. The nanosized structure enables the uptake of OMVs by antigen presenting cells, and facilitates efficient transport through lymphatic system [34]. It is highly conceivable that similar OMV-based vaccines could be developed against *F. nucleatum* infection. While *F. nucleatum* OMVs have been shown to contain various antigenic elements, a comprehensive analysis of all the components found in OMVs generated by *F. nucleatum* has yet to be completed.

In this study, we isolated and purified OMVs from *F. nucleatum subsp animalis 7_1 (EAVG_002)* and biophysically and proteomically profiled the purified OMVs to explore potential development strategies for OMV-based *F. nucleatum* vaccines. The diameter and shape of OMVs were measured using transmission electron microscopy and dynamic light scattering. OMVs' constituents and membrane fraction composition were analyzed by nano LC/MS/MS and characterized by bioinformatics tools. The results from our comprehensive analysis of *F. nucleatum* OMVs as reported below can serve as a firm foundation for further discoveries regarding their role in pathogenesis and anti-*F. nucleatum* vaccine development.

Materials and Methods

Bacterial strains and cultivation

F. nucleatum subsp animalis 7_1 (EAVG_002) was kindly provided by Dr. Emma Allen-Vercoe (University of Guelph) [35]. The strain was grown anaerobically (80% N_2, 10%CO_2, 10% H_2) at 37°C either on Fastidious Anaerobe Agar (Neogen) with 5% horse blood or in tryptic soy broth (Becton, Dickinson and Company) with 5 μg/mL hemin and 5 μg/mL menadione (Sigma-Aldrich)[35]. The bacteria were harvested at OD_{600} 0.7~0.9. A volume of each strain was mixed with an equal volume of glycerol and stored at -80 °C in cryogenic storage tubes for further use.

OMVs isolation and purification

The OMVs of *F. nucleatum* were collected from the supernatant of bacterial culture (200mL) according to the method described by Chutkan et al. with modifications[36-37]. The process flow chart is shown in Figure 2. Briefly, cells were spun down by centrifugation at 8500 x g for 15 min at 4°C (rotor JA-14, Beckman-Coulter), then the supernatant was vacuum filtered through a 0.22-

μm membrane to remove any remaining bacteria. The OMVs were collected by ultracentrifuge at 213,000 x g for 2hr at 4°C (rotor Type 50.2 Ti, Beckman-Coulter). The pelleted OMVs were suspended with sterile phosphate-buffered saline (PBS). The resuspended OMVs were purified by density gradient centrifugation with OptiPrep™ (60% iodixanl (w/v), Axis-Shield). The OptiPrep™ solution was diluted with 0.85% (w/v) NaCl containing 10mM Tricine-NaOH pH 7.4 into 35%, 30%, 25%, 20% (v/v) OptiPrep™. The 40% OptiPrep™ solution was prepared with suspended OMVs and layered 2mL at the bottom of an Ultra-Clear™ centrifuge tube (13.2 mL, Beckman Coulter). The discontinuous gradients were loaded in 2 mL increments in degressive concentrations on top of the prior layer [36]. Prepared tubes were centrifuged at 135,000x g for 16 hr at 4°C (rotor SW 41 Ti, Beckman-Coulter). Each 1 mL fraction was carefully collected by pipet from top to bottom. A portion of each fraction was analyzed by 12% SDS-page, which displayed that the OMVs were mainly contained in fraction 25%~30% (data not shown). The OMV-containing fractions were collected and pooled into an ultracentrifugation tube with 27 mL sterile PBS (at least 10-fold the sample volume) and ultracentrifuged at 200,500 x g for 2hr at 4°C (rotor Type 45 Ti, Beckman Coulter) to remove the OptiPrep™ solution. The purified OMVs were resuspended with sterile PBS and the aliquots were stored at -80°C for future use.

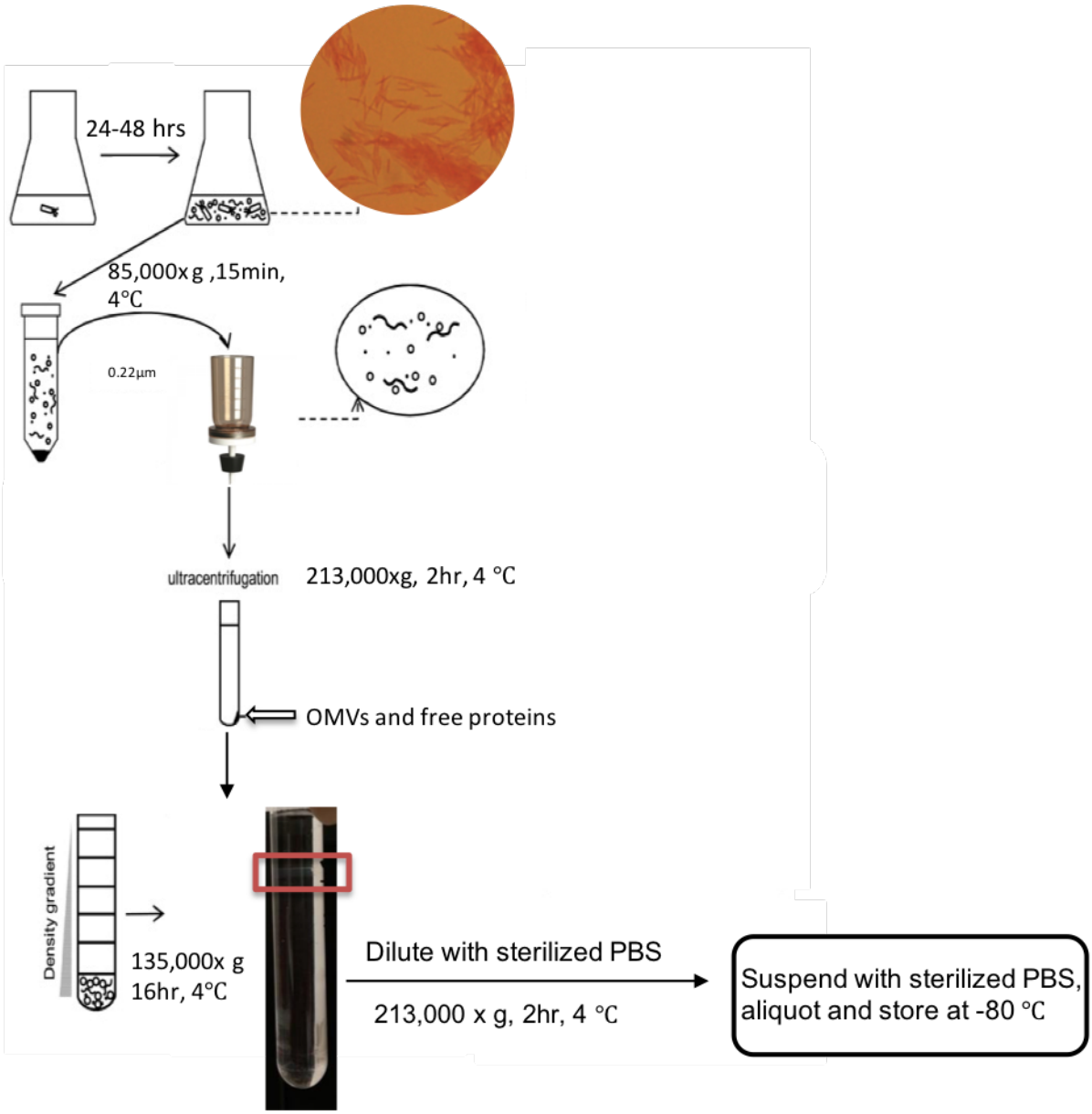

24-48 hrs
85,000x g ,15min,
4°C
0.22μm
ultracentrifugation
213,000xg, 2hr, 4 °C
OMVs and free proteins
Density gradient
135,000x g
16hr, 4°C
Dilute with sterilized PBS
213,000 x g, 2hr, 4 °C
Suspend with sterilized PBS,
aliquot and store at -80 °C

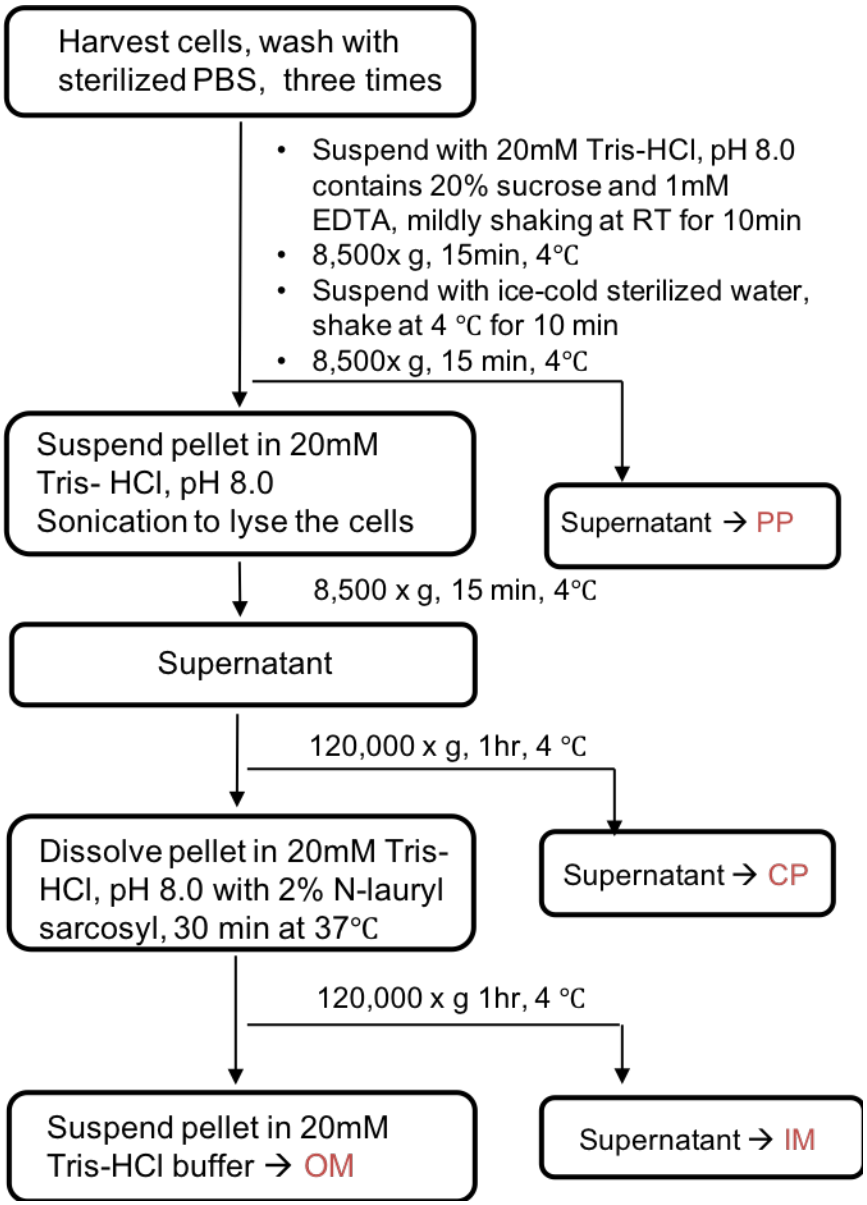
Figure 2 Flow chart of OMVs isolation and purification

The overnight-cultured bacteria are spun down at low speed to exclude the whole bacteria. The supernatant is filtrated with 0.22 μm membrane to remove the debris. The filtrated media is ultracentrifuged at 213,000 x g for 2 hours 4 °C. The pellet containing OMVs is resuspend with PBS and further purified by gradient purification at 135,000 x g for 16 hours at 4 °C. The OMVs can be visualized and further collected by dilute with PBS and ultracentrifuge at 200,500 x g for 2 hours 4 °C. The product was aliquoted and stored at – 80 °C.

The figure is adapted from Klimentová et al. Microbiological Research (2015) [37].

Cellular fraction preparation

The cellular fractions were obtained from the same culture as the OMV preparations, based on the protocol by Aguilera et al. with modifications [38]. The harvested cells were washed with sterile PBS three times and suspended in 20 mM Tris–HCl, pH 8.0 with 20% sucrose and 1mM EDTA, then mildly shaken at room temperature for 10 min. The solution was spun at 8500 x g for 15 min at 4°C (rotor JA-14, Beckman-Coulter), immediately suspended in ice-cold sterile water (200mL), and gently shaken at 4°C for 10 min. The supernatant with periplasmic components was collected by centrifugation (8500 x g for 15 min at 4°C). The cell pellets were suspended in 20 mM Tris-HCl, pH 8.0 (25mL) and disrupted by sonication. The suspension was centrifuged at 8500 x g for 15 min at 4°C; the supernatant was ultracentrifuged at 120,000 x g for 1 hr at 4°C. The pellet was dissolved in 20 mM Tris–HCl, pH 8.0 containing 2% N-lauryl sarcosyl (Sigma-Aldrich) (5 mL) for 30 min at 37ºC. The supernatant consisting of inner membrane portion and the pellet contained outer membrane proteins was suspended in 20 mM Tris–HCl buffer (1 mL). The periplasmic fraction was (v/v) concentrated 100:1 (Millipore) and all fractions were filtered using a 0.45-µm PVDF membrane (Millipore) to remove remaining cells. A portion (12.5 µL) of each sample was analyzed by 12% SDS-PAGE and the rest was stored at -80°C.



*PP- periplasmic fraction
CP- cytoplasmic fraction
IM – inner membrane fraction
OM- outer membrane fraction

Figure 3 The flow chart of prepare cellular fractions

The harvested cells were washed three times with sterile PBS and suspended in 20 mM Tris–HCl, pH 8.0 with 20% sucrose and 1mM EDTA. After mildly shaken at room temperature for 10 min, the solution was spun at 8500 x g for 15 min at 4°C (rotor JA-14, Beckman-Coulter). Then immediately suspended the pellet in ice-cold sterile water, and gently shaken at 4°C for 10 min. The supernatant with periplasmic components was collected by centrifugation (8500 x g for 15 min at 4°C). The cell pellets were suspended in 20 mM Tris-HCl, pH 8.0 (25mL) and disrupted by sonication. The suspension was centrifuged at 8500 x g for 15 min at 4°C to remove the whole cells and debris; the supernatant was ultracentrifuged at 120,000 x g for 1 hr at 4°C. The pellet was dissolved in 20 mM Tris–HCl, pH 8.0 containing 2% N-lauryl sarcosyl (Sigma-Aldrich) for 30 min at 37ºC. The supernatant consisting of inner membrane portion and the pellet contained outer membrane proteins was suspended in 20 mM Tris–HCl buffer.

Dynamic light scattering (DLS)

The size distribution of OMVs in sterile PBS was determined with DLS using a Malvern Nano ZS Zetasizer (Malvern Panalytical). Triplicate measurements of 100 ng/μL OMVs samples in 1.0 mL were loaded into disposable sizing cuvettes at 25°C, and the system was assigned with 1.330 dispersant refractive index and 0.8882 viscosity. The average diameter (Z-average) of OMVs were collected and analyzed with the Zetasizer software.

Transmission electron microscopy (TEM)

Samples were negative stained with 1.5% aqueous Uranyl Acetate on 300 mesh carbon coated copper TEM grids. The images were acquired using a Thermo-Fisher Tecnai Spirit Twin STEM equipped with a SIS MegaView III camera in TEM mode at 120kV.

In-gel trypsin digestion of SDS gel bands

The protein bands from a 12% SDS-PAGE gel (2 to 4 gel slices for each of the 6 samples) were cut into ~1 mm slices and subjected to in-gel digestion followed by extraction of the tryptic peptide as reported previously [39]. The excised gel pieces were washed consecutively in 200 μL distilled water, 100 mM ammonium bicarbonate (Ambic)/acetonitrile (1:1) and acetonitrile. The gel pieces were reduced with 70 μL of 10 mM DTT in 100 mM Ambic for 1 hr at 56°C, alkylated with 100 μL of 55 mM Iodoacetamide in 100 mM Ambic at room temperature in the dark for 1 hr. After washing gel slices as described above, the gel slices were dried and rehydrated with 50 μL trypsin (Promega, Madison, WI) in 50 mM Ambic, 10% ACN (20 ng/μL) at 37°C for 16 hrs. The digested peptides were extracted twice with 70 μl of 50% acetonitrile, 5% FA and once with 70 μl of 90% acetonitrile, 5% FA. Extracts from each sample were combined and lyophilized.

Protein Identification by nano LC/MS/MS Analysis

In-gel tryptic digests were reconstituted in 20 µL of 0.5% FA for nano LC-ESI-MS/MS analysis, which was carried out using an Orbitrap Fusion™ Tribrid™ (Thermo-Fisher Scientific, San Jose, CA) mass spectrometer equipped with a nanospray Flex Ion Source, and coupled with a Dionex UltiMate3000RSLCnano system (Thermo, Sunnyvale, CA) [40-41]. Gel extracted peptide samples (5-15 µL) were injected onto a PepMap C-18 RP nano trapping column (5 µm, 100 µm i.d x 20 mm) at 20 µL/min flow rate for rapid sample loading and then separated on a PepMap C-18 RP nano column (2 µm, 75 µm x 25 cm) at 35°C. The tryptic peptides were eluted in a 120 min gradient of 5% to 38% acetonitrile (ACN) in 0.1% formic acid solution at 300 nL/min., followed by a 7 min ramping to 90% ACN-0.1% FA and an 8 min hold at 90% ACN-0.1% FA. The column was re-equilibrated with 0.1% FA for 25 min prior to the next run. The Orbitrap Fusion was operated in positive ion mode with spray voltage set at 1.6 kV and source temperature at 275°C. External calibrations for FT, IT and quadrupole mass analyzers were performed. In data-dependent acquisition (DDA) analysis, the instrument was operated using FT mass analyzer in MS scan to select precursor ions followed by 3-second "Top Speed" data-dependent CID ion trap MS/MS scans at 1.6 m/z quadrupole isolation for precursor peptides with multiple charged ions above a threshold ion count of 10,000 and normalized collision energy of 30%. MS survey scans at a resolving power of 120,000 (fwhm at m/z 200) for the mass range of m/z 375-1575. Dynamic exclusion parameters were set at 40 s of exclusion duration with ±10 ppm exclusion mass width. All data were acquired under Xcalibur 3.0 operation software (Thermo-Fisher Scientific).

Data analysis

The DDA raw files for CID MS/MS were subjected to database searches using Proteome Discoverer (PD) 2.2 software (Thermo Fisher Scientific, Bremen, Germany) with the Sequest HT algorithm. The PD 2.2 processing workflow containing an additional node of Minora Feature Detector for precursor ion-based quantification was used for protein identification and relative quantitation analysis. The database search was conducted against a combined database containing *Fusobacterium nucleatum* RefSeq database 2416 entries downloaded from UniProt. Two-missed trypsin cleavage sites were allowed. The peptide precursor tolerance was set to 10 ppm and fragment ion tolerance was set to 0.6 Da. Variable modification of methionine oxidation, N-terminal acetylation and deamidation of asparagine/glutamine and fixed modification of cysteine carbamidomethylation, were set for the database search. Only high confidence peptides defined by Sequest HT with a 1% FDR by Percolator were considered for the peptide identification. The final protein IDs contained protein groups that were filtered with at least 2 peptides per protein.

Bioinformatics analysis

Sequence of each protein was acquired from both UniProt database (http://www.uniprot.org/) of *F. nucleatum* subsp. animalis 7-1 and NCBI (*F. nucleatum* RefSeq20170725) Protein database (https://www.ncbi.nlm.nih.gov/protein/) [42-43]. Outer membrane protein beta-barrels were predicted by PRED-TMBB2 (http://www.compgen.org/tools/PRED-TMBB2) [44]. Subcellular localization, adhesion affinity levels, and transmembrane helices of proteins were obtained from Vaxign (http://www.violinet.org/vaxign/) [45-49]. Protein localization was predicted by Cello2go (http://cello.life.nctu.edu.tw/cello2go/), and signal peptides, including lipoproteins, were predicted by LipoP 1.0 Server (http://www.cbs.dtu.dk/services/LipoP/) [50-51]. Putative function domains of

each protein was generated from InterPro (https://www.ebi.ac.uk/interpro/) [52]. Non-classical protein secretion was anticipated by SecretomeP 2.0 Server (http://www.cbs.dtu.dk/services/SecretomeP/) for proteins without predicted signal peptide [53]. All of the proteins with predicted signal peptides contain exposed antigen peptides predicted by EMBOSS (http://www.bioinformatics.nl/cgi-bin/emboss/antigenic) with EMBOSS Motif output report format and six minimum lengths of regions for the various antigenic sites [54-56]. Antigenic B-cell epitopes were predicted using BCPREDS: B-cell epitope prediction server (http://ailab.ist.psu.edu/bcpred/predict.html) [57]. In this application, two methods-BCPred (fixed length epitope prediction) and FBCPred (Flexible length epitope prediction) with assigned length 12, 75% specificity and a prediction score larger than 0.9 were selected and shown in the table. To identify T-cell epitopes Vaxitop (http://www.violinet.org/vaxign/vaxitop/index.php) was applied with P-value cut-off 0.05 for all HLA-A2 and HLA-A3 alleles; the classes selection were based on previous publications that they showed overexpressed patterns in colon tumor tissues [45-46, 58-59]. Metabolic pathways were classified with KEGG pathway database (http://www.genome.jp/kegg/pathway.html) [60]. Functional groups were classified by EggNOG 4.5.1 (http://eggnogdb.embl.de/#/app/home) [61]. Cargo sorting to the OMV in *F.nucleatum* was plotted by ggplot2 package with R studio [62].

Results

Production and purity of OMVs from *F. nucleatum* subsp animalis 7_1 (EAVG_002)

The OMVs collected from the supernatant of *F. nucleatum subsp animalis 7_1 (EAVG_002)* broth were verified by transmission electron microscopy (Figure 4) and dynamic light scattering (Figure 5). The OMVs appear varied ranging in size from 40 to 110 nm with spherical shapes; the typical

diameter was around 79 nm (Figure 5). Duplicate OMV preparations and cellular fractions from *F. nucleatum* were separated on an SDS-PAGE gel. Each gel-lane was excised into multiple sections (Figure 6) and analyzed by mass spectrometer. Among these, 157 proteins were identified from the first duplication and 190 proteins were identified from the second duplication, with 98 overlapping proteins identified in both duplications (Figure 7, Table 2). Purity of OMVs was estimated by comparing these to abundant proteins in the cytosol and inner membrane fractions. Abundant cytosol proteins were assessed by high number of peptide spectrum matches (PSMs) and by those lacking predicted signal peptides. Among the top 30 cytosol proteins with highest PSMs, only one (A0A140PU88) was identified as one of the 98 OMV proteins (Figure 8). Five of the top 30 inner membrane proteins were also recognized on OMVs, which are WP_008700001.1, WP_008701980.1, WP_016361171.1, A0A140PRA0, and WP_008701946.1 (Figure 9). All five proteins were detected in the periplasmic and outer membrane fractions, suggesting these proteins may break free from the inner membrane, migrate to other fractions and ultimately bind to OMVs (Figure 9). Three of these five proteins were classified as autotransporters, which might explain their appearance in multiple fractions. The near absence of cytoplasmic proteins in OMVs implies that the level of debris contamination was negligible.

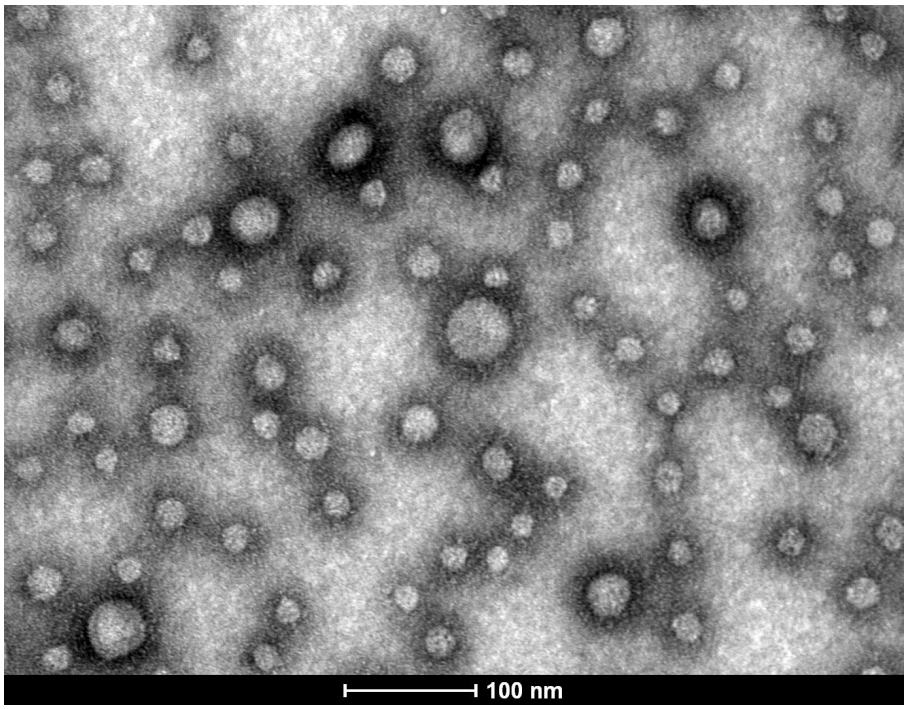


Figure 4 TEM image of *F.nucleatum* OMVs

A TEM image of OMVs in size varied from 40 to 100 nm with spherical shape. Majority has the size around 40.

			Size (d.nm):	% Intensity:	St Dev (d.n...
Z-Average (d.nm):	78.74	Peak 1:	94.38	93.6	54.47
PdI:	0.244	Peak 2:	1015	4.8	504.2
Intercept:	0.970	Peak 3:	3753	1.6	1065
Result quality :	Good				

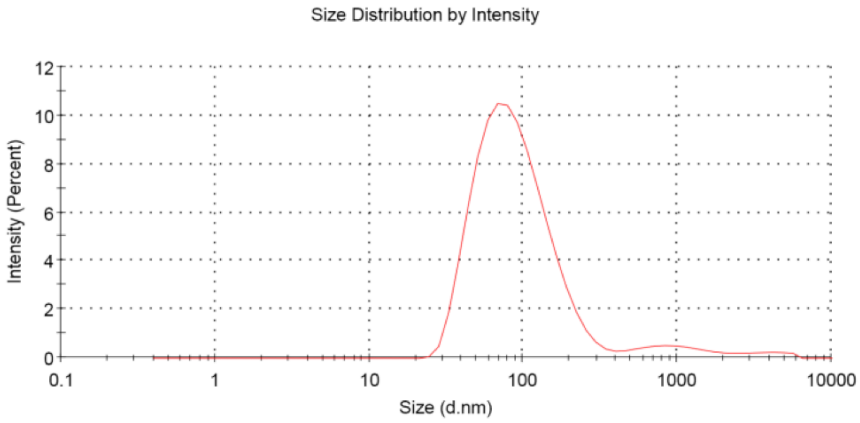


Figure 5 The DLS measurement of OMV

The average size of OMV is 78. 74 nm, 93.6% intensity is identified at size 94.38 nm. The measurement is with good quality.

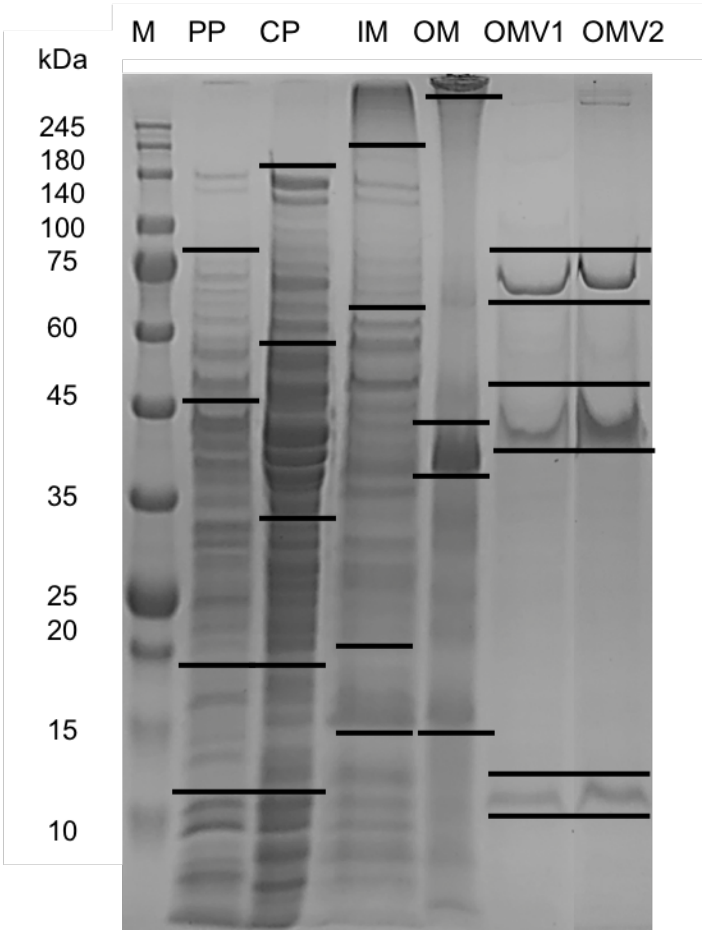


Figure 6 SDS-PAGE gel excision

SDS-PAGE of *F.nucleatum* fractions and OMVs used for proteome analysis and duplicated identified proteins from OMVs. SDS-PAGE of subcellular fraction for periplasm (PP), cytoplasm (CP), inner membrane (IM), outer membrane (OM) and duplication of OMV1&2. The gel excision was as shown.

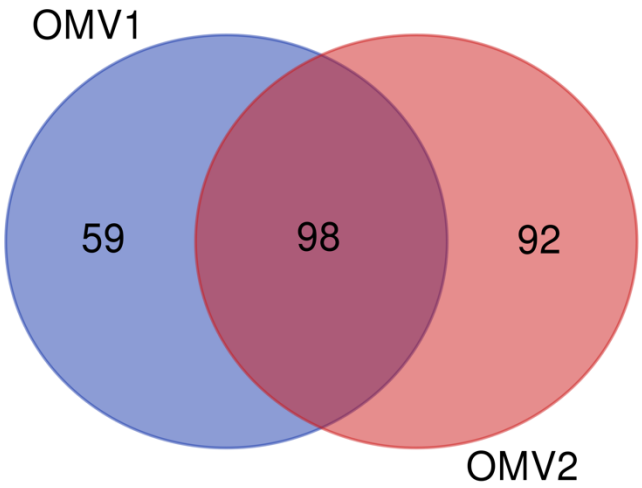


Figure 7 Venn diagram of the duplication OMVs

98 overlapping proteins were identified from the two batches of OMVs.

Table 2 98 Identified Proteins on OMVs

Accession	Description	Pfam IDs	MW [kDa]	calc. pI	ave mol% OMV[a]	mol% OMV[b]	Beta-barrel	Adhesin[d]	Class	Localization
A0A140PNL7	Uncharacterized protein	Pf00691, Pf13488	21.8	9.51	0.565847674	0.297330204	N	0.052	0.458	Lipoprotein	Outer Membrane
A0A140PP62	60 kDa chaperonin	Pf00118	57.6	5.03	0.005414283	0	N	0.008	0.18	Cyt	Cytoplasmic
A0A140PP79	30S ribosomal protein S13	Pf00416	13.6	10.35	0.132485831	0.005442428	N	0.004	0.218	Cyt	Cytoplasmic
A0A140PP98	Uncharacterized protein	Pf00691	20.2	8.82	4.054411368	0.825160564	N	0.023	0.147	Lipoprotein	Outer Membrane
A0A140PPN2	Uncharacterized protein	Pf03480, Pf13654	38.9	5.27	0.051481799	0.139904097	N	0.004	0.376	Lipoprotein	Unknown
A0A140PPN3	Uncharacterized protein	Pf10634	24.6	4.72	0.114169735	0.350236311	N	0.038	0.33	Lipoprotein	Unknown
A0A140PPW6	50S ribosomal protein L13	Pf00572	16.4	9.86	0.953620651	0.001749565	N	0.000	0.176	Cyt	Cytoplasmic
A0A140PQA9	Uncharacterized protein	Pf00496, Pf00497	56.1	5.26	0.063351671	0.017856288	N	0.006	0.149	Lipoprotein	Periplasmic
A0A140PQL8	30S ribosomal protein S16	Pf00886	9.9	10.1	0.072302955	0.00283966	N	0.000	0.081	Cyt	Cytoplasmic
A0A140PQR2*	Uncharacterized protein	Pf03864	37.1	5.92	0.010505526	0.886254676	N	0.001	0.449	Cyt	Cytoplasmic
A0A140PQV3	Uncharacterized protein		18.7	4.53	0.001197804	0.040562312	N	0.066	0.507	Lipoprotein	Unknown
A0A140PRA0	Uncharacterized protein	Pf00441, Pf02770, Pf02771, Pf08028	41.6	7.24	0	0.085125474	N	0.000	0.311	Cyt	Cytoplasmic
A0A140PRH8	Uncharacterized protein	Pf00216	10.4	9.94	0.088089699	0.004081367	N	0.000	0.161	Cyt	Cytoplasmic
A0A140PRZ6	Ribosome hibernation promoting factor	Pf02482	20.6	9.26	0.114705841	0.02619233	N	0.001	0.141	Cyt	Cytoplasmic
A0A140PS00#	Uncharacterized protein	Pf09403	14.4	4.88	0.023970582	0.002219003	Y	0.704	0.475	Spl	Unknown
A0A140PSC3	Phosphate/phosphite/phosphonate ABC transporters, periplasmic binding protein	Pf00497, Pf09084, Pf12974, Pf13531	32	6.01	0.100603491	0.209824047	N	0.044	0.329	Lipoprotein	Periplasmic
A0A140PSJ8	Uncharacterized protein	Pf09864	13.4	4.51	2.458609358	0	N	0.050	0.511	Lipoprotein	Periplasmic
A0A140PSK0	Uncharacterized protein		16.2	9.14	6.62118053	0.219059389	Y	0.467	0.522	Lipoprotein	Unknown
A0A140PSK8	Uncharacterized protein		28.6	5.27	0.030300311	0.09781311	N	0.034	0.493	Lipoprotein	Unknown
A0A140PSQ8	Uncharacterized protein		6.3	9.47	2.632485716	0.061069395	N	0.000	0.446	Lipoprotein	Unknown
A0A140PSZ8	Acetyl-CoA C-acetyltransferase	Pf00108, Pf00109, Pf02803	42.4	7.87	0	0.79829135	N	0.004	0.214	Cyt	Cytoplasmic
A0A140PTC5	30S ribosomal protein S8	Pf00410	14.8	9.91	0.379796853	0.051594779	N	0.003	0.14	Cyt	Cytoplasmic
A0A140PTM6	50S ribosomal protein L5	Pf00281, Pf00673	20.7	9.28	0.200096825	0.037547612	N	0.003	0.07	Cyt	Cytoplasmic
A0A140PTN6	Uncharacterized protein	Pf01297	34.1	9.16	0.016792754	0.004137767	N	0.002	0.206	Spl	Periplasmic
A0A140PTQ0#	Uncharacterized protein	Pf03724	16.3	8.92	0.241424367	0.012678916	N	0.005	0.395	Lipoprotein	Unknown
A0A140PTY8*	30S ribosomal protein S20	Pf01649	9.9	11	0.096763137	0.001235261	N	0.000	0.379	Cyt	Cytoplasmic
A0A140PU12	Uncharacterized protein	Pf13460	36.8	6.1	0.077892463	0.025542952	N	0.004	0.127	Cyt	Cytoplasmic
A0A140PU88	Elongation factor Tu	Pf00009, Pf03143, Pf03144	43.5	5.38	0.120592455	0.027182808	N	0.001	0.11	Cyt	Cytoplasmic
A0A140PU90	Uncharacterized protein	Pf03780	13.4	4.55	0	0.008092285	N	0.039	0.05	Cyt	Cytoplasmic
A0A140PUD9	Uncharacterized protein		42.4	8.88	23.12004172	39.69290988	Y	1.000	0.633	Spl	Outer Membrane
A0A140PUE5	Uncharacterized protein		6.7	9.13	0.001392975	0	N	0.000	0.065	Cyt	Unknown
A0A140PUN8#	Uncharacterized protein	Pf10670	28.8	8.24	0.042164836	0.018489233	N	0.003	0.393	Spl	Unknown
A0A140PUR5	Uncharacterized protein		14.6	5.12	0.127287034	0.156596281	N	0.077	0.251	Lipoprotein	Unknown
A0A140PUV5	30S ribosomal protein S10	Pf00338	11.5	10.05	0.119575652	0.016154558	N	0.000	0.147	Cyt	Cytoplasmic

Accession	Description	Pfam IDs	MW [kDa]	calc. pI	ave mol% OMV[a]	mol% OMV[b]	Beta-barrel	Adhesin[d]	Class	Localization
A0A140PUZ8	Uncharacterized protein		19.1	8.05	0.026013447	0.000272439	N|0.018	0.395	Lipoprotein	Unknown
A0A140PVG6	50S ribosomal protein L22	Pf00237	12.3	10.56	0.029094535	0.009863248	N|0.000	0.092	Cyt	Cytoplasmic
A0A140PVH3	50S ribosomal protein L6	Pf00347	19.4	9.83	0.127263018	0.011294608	N|0.002	0.188	Cyt	Cytoplasmic
A0A140PVN2	Uncharacterized protein		16.1	9.45	0.012449469	0.175659455	N|0.030	0.545	SpI	Unknown
A0A140PVZ1	50S ribosomal protein L10	Pf00466	18.9	9.19	0.14822238	0.011344908	N|0.018	0.356	Cyt	Cytoplasmic
A0A140PW10	30S ribosomal protein S2	Pf00318	27.9	5.95	0.019618097	0.043882652	N|0.003	0.15	Cyt	Cytoplasmic
A0A140PW34	50S ribosomal protein L11	Pf00298, Pf03946	14.9	9.6	0.000835167	0.007170161	N|0.016	0.436	Cyt	Cytoplasmic
A0A140PWC8	Uncharacterized protein		15.9	8.98	0.163939687	0.019910152	Y|0.580	0.462	Lipoprotein	Unknown
A0A140PWZ2	50S ribosomal protein L4	Pf00573	22.7	9.82	0.00113116	0.01936733	N|0.016	0.249	Cyt	Cytoplasmic
NP_602362.1	acyl-CoA dehydrogenase	Pf00441, Pf02770, Pf02771, Pf08028	41.2	5.3	0.035929299	0	N|0.000	0.342	Cyt	Cytoplasmic
NP_602384.1*	30S ribosomal protein S7	Pf00177	17.8	9.77	0.024203183	0.013933933	N|0.004	0.088	Cyt	Cytoplasmic
NP_602455.1*	30S ribosomal protein S3	Pf00189, Pf07650	24.3	9.76	0.040057916	0.014672749	N|0.003	0.292	Cyt	Cytoplasmic
NP_602472.1*[+]	30S ribosomal protein S18	Pf01084	8.4	11.25	0.320034119	0	N|0.000	0.041	Cyt	Cytoplasmic
NP_602706.1[+]	Cytosol aminopeptidase	Pf00883, Pf02789	53.6	8.92	0	0	N|0.001	0.285	Cyt	Cytoplasmic
NP_602821.1	DNA-directed RNA polymerase subunit beta'	Pf00529, Pf00623, Pf04983, Pf04997, Pf04998, Pf05000	148	8.78	0.029243508	0	N|0.000	0.05	Cyt	Cytoplasmic
NP_603579.1	fibronectin-binding protein-like protein A	Pf05670, Pf05833	63	8.94	0	0	N|0.001	0.265	Cyt	Cytoplasmic
NP_604327.1	short chain dehydrogenase	Pf00106, Pf08659, Pf13561	28.9	6.61	0.010764492	0.000954424	N|0.018	0.142	Cyt	Cytoplasmic
WP_005895855.1	aspartate--tRNA ligase	Pf00152, Pf00587, Pf01336, Pf02938	67.5	5.45	0.015982857	0	N|0.002	0.054	Cyt	Cytoplasmic
WP_005899012.1	phosphate acetyltransferase	Pf01515	36	7.69	0.005656867	0	N|0.006	0.159	Cyt	Cytoplasmic
WP_005902041.1	type I glyceraldehyde-3-phosphate dehydrogenase	Pf00044, Pf02800, Pf03435	35.9	6.01	0.00116944	0	N|0.006	0.254	Cyt	Cytoplasmic
WP_005906342.1	hypothetical protein		28.9	9.04	0.614204242	0.115087315	N|0.002	0.386	Lipoprotein	Cytoplasmic
WP_005908723.1	hypothetical protein		33.7	9.29	0.2144239	0.060685723	N|0.013	0.467	Lipoprotein	Cytoplasmic
WP_005908903.1	DUF5105 domain-containing protein		19.2	5.15	0.008980763	0.007110022	N|0.005	0.411	Lipoprotein	Unknown
WP_005910001.1[#]	DUF3798 domain-containing protein	Pf12683, Pf13377	45.6	5.1	0.057312307	0	N|0.008	0.35	Lipoprotein	Unknown
WP_005910189.1*	peptidylprolyl isomerase	Pf00160	30.8	7.97	0.065417964	0.127271884	N|0.014	0.584	Lipoprotein	Cytoplasmic
WP_005910368.1*	hypothetical protein	Pf03961, Pf09403	15.9	5.21	1.496198965	0.306714667	N|0.013	0.34	SpI	Unknown
WP_005910677.1	tetratricopeptide repeat protein	Pf00515, Pf07719, Pf13176, Pf13181, Pf13374, Pf13414, Pf13424, Pf13432, Pf14559	31.5	9.07	0.00120992	0	N|0.022	0.331	Lipoprotein	Cytoplasmic
WP_005910688.1*	hypothetical protein		10.8	5.34	0.016596106	0.002345745	N|0.000	0.507	Lipoprotein	Unknown
WP_008691343.1	hypothetical protein		47.2	8.95	0.15142948	0.085961248	N|0.012	0.485	Lipoprotein	Outer Membrane
WP_008691364.1	hypothetical protein		15.3	9.22	0.002204904	0	N|0.268	0.290	SpI	Unknown

Accession	Description	Pfam IDs	MW [kDa]	calc. pI	ave mol% OMV[a]	mol% OMV[b]	Beta-barrel	Adhesin[d]	Class	Localization
WP_008692599.1[#]	toxin-antitoxin system YwqK family antitoxin		29.3	8.79	0.080650357	0	N|0.111	0.307	SpI	Unknown
WP_008692881.1	iron(III)-binding protein	Pf01547, Pf12849, Pf13343, Pf13416, Pf13531	38.8	4.88	0.031824278	0.369821817	N|0.002	0.329	Lipoprotein	Periplasmic
WP_008693244.1[+]	50S ribosomal protein L24		12.5	10.24	0.652254427	0.131412403	N|0.000	0.208	Cyt	Cytoplasmic
WP_008693392.1	glycerophosphoryl diester phosphodiesterase	Pf03009	41.3	8	0.001813713	0.003413634	N|0.013	0.539	SpI	Unknown
WP_008694231.1	N-acetylmuramoyl-L-alanine amidase	Pf01471, Pf01510, Pf08139	34.4	7.58	0.100439494	0	N|0.006	0.21	Lipoprotein	Cytoplasmic
WP_008694296.1*	hypothetical protein		13.8	4.49	0.042863871	0	N|0.079	0.526	SpI	Unknown
WP_008694670.1	ATP-dependent Clp protease ATP-binding subunit ClpX	Pf00004, Pf01078, Pf04851, Pf05496, Pf06689, Pf07724, Pf07726, Pf07728, Pf10431, Pf13191, Pf13207, Pf13401	47.5	7.27	0.01154732	0	N|0.001	0.21	Cyt	Cytoplasmic
WP_008700001.1*	diguanylate phosphodiesterase	Pf00496	57.1	5.36	1.592780482	0.476946236	N|0.008	0.274	Lipoprotein	Periplasmic
WP_008700150.1*	hypothetical protein		54.7	8.87	0.05998915	0.005252595	N|0.037	0.433	Lipoprotein	Outer Membrane
WP_008700191.1*	hypothetical protein		9.7	4.96	0.051431224	1.387152671	N|0.000	0.452	Cyt	Unknown
WP_008700266.1*	cell surface protein	Pf03895, Pf05662	75.7	8.06	0.181236059	0.397920868	Y|1.000	0.482	SpI	Outer Membrane
WP_008700372.1*	cell surface protein	Pf03895, Pf04977, Pf05662, Pf12329	54.3	9.45	0.017332556	0.043146486	Y|1.000	0.642	SpI	Extracellular
WP_008701093.1	complement resistance protein TraT	Pf05818	25.6	9.16	0.098981677	0.272315164	N|0.012	0.521	Lipoprotein	Outer Membrane
WP_008701556.1*	autotransporter-associated N-terminal domain-containing protein	Pf03797	232.2	9.1	0.104297369	2.834379118	Y|1.000	0.572	SpI	Outer Membrane
WP_008701805.1*[#]	membrane protein	Pf03938	17.7	9.06	0.098855986	0	N|0.007	0.604	SpI	Unknown
WP_008701946.1*	autotransporter domain-containing protein	Pf03797	255.1	8.78	0.186175293	3.600222526	Y|1.000	0.63	SpI	Unknown
WP_008701980.1*	autotransporter-associated N-terminal domain-containing protein	Pf03797	368.2	8.56	3.414931136	3.421590985	Y|0.992	0.726	SpI	Unknown
WP_008702357.1*	acetyl-CoA--acetoacetyl-CoA transferase subunit alpha	Pf01144, Pf02550	23.2	5.91	0	0.032350141	N|0.004	0.11	Cyt	Cytoplasmic
WP_008702403.1*[#]	hypothetical protein	Pf13424	109.8	8.72	1.059710764	0	N|0.002	0.441	SpI	Outer Membrane
WP_008702680.1*[#]	toxin-antitoxin system YwqK family antitoxin		27.2	8.72	0.047517769	0	N|0.066	0.534	SpI	Unknown
WP_008795754.1	DUF5105 domain-containing protein		19.2	5.73	0.095715319	0	N|0.018	0.457	Lipoprotein	Unknown

Accession	Description	Pfam IDs	MW [kDa]	calc. pI	ave mol% OMV[a]	mol% OMV[b]	Beta-barrel		Adhesin[d]	Class	Localization
WP_008796203.1[#]	hypothetical protein		23.9	9.48	0.032063344	0	N	0.076	0.671	SpI	Unknown
WP_016339730.1	hypothetical protein		21	8.16	0.01145081	0	N	0.321	0.394	SpI	Unknown
WP_016361140.1*[#]	flavoCytochrome c	Pf00890, Pf04205, Pf12831	60.7	9.35	1.002747394	0	N	0.001	0.444	SpI	Periplasmic
WP_016361171.1*	autotransporter-associated beta strand	Pf00082, Pf03797	118.1	7.27	36.80949397	5.667577584	Y	1.000	0.546	Lipoprotein	Outer Membrane
WP_016361305.1*	hypothetical protein		12.3	9.23	0.118763424	0	N	0.000	0.269	Lipoprotein	Unknown
WP_016361319.1*[#]	hypothetical protein	Pf03349	53.5	9.13	2.601973906	0.7176589	Y	1.000	0.611	SpI	Outer Membrane
WP_016361366.1*	autotransporter domain-containing protein	Pf03797	261.3	8.9	0.58748604	2.562998039	Y	1.000	0.645	Lipoprotein	Unknown
WP_017139894.1	hypothetical protein	Pf03895	164.7	9.6	0.035019081	0.103059819	Y	0.916	0.681	SpI	Outer Membrane
WP_020788899.1	autotransporter domain-containing protein	Pf03797	225.1	8.84	0.003160904	0	Y	1.000	0.619	SpI	Unknown
WP_023038391.1	outer membrane protein assembly factor	Pf01103, Pf07244, Pf08478	78.3	8.25	0.462906494	0	Y	1.000	0.43	SpI	Outer Membrane
WP_029758176.1	hypothetical protein	Pf00092, Pf13768	27.6	6.34	0.338219251	0.025141289	N	0.008	0.49	SpI	Unknown
WP_060675827.1	hypothetical protein	Pf00691	21.6	6.24	0.087976239	0	N	0.020	0.2	SpI	Outer Membrane
WP_060676683.1	hypothetical protein		42.9	8.88	0	0	Y	1.000	0.604	SpI	Outer Membrane

Table 2. 98 identified proteins in both OMV replications.

[a] Average molar % (by abundance) of each protein calculated from replicated OMV samples. [b] Molar% of outer membrane fraction for identified protein. [c] Prediction of transmembrane segments by PRED-TMBB with cut-off 0.43. [d] Predicted adhesion probability of each protein. * The protein had been identified both from UniProt and NCBI databases. [#] The protein had been identified both on outer membrane and periplasm fraction with a molar% ratio smaller than 0.4. [+]The proteins classified as cytosol protein and without signal peptide were predicted as non-classically secreted protein by recommend threshold 0.5.

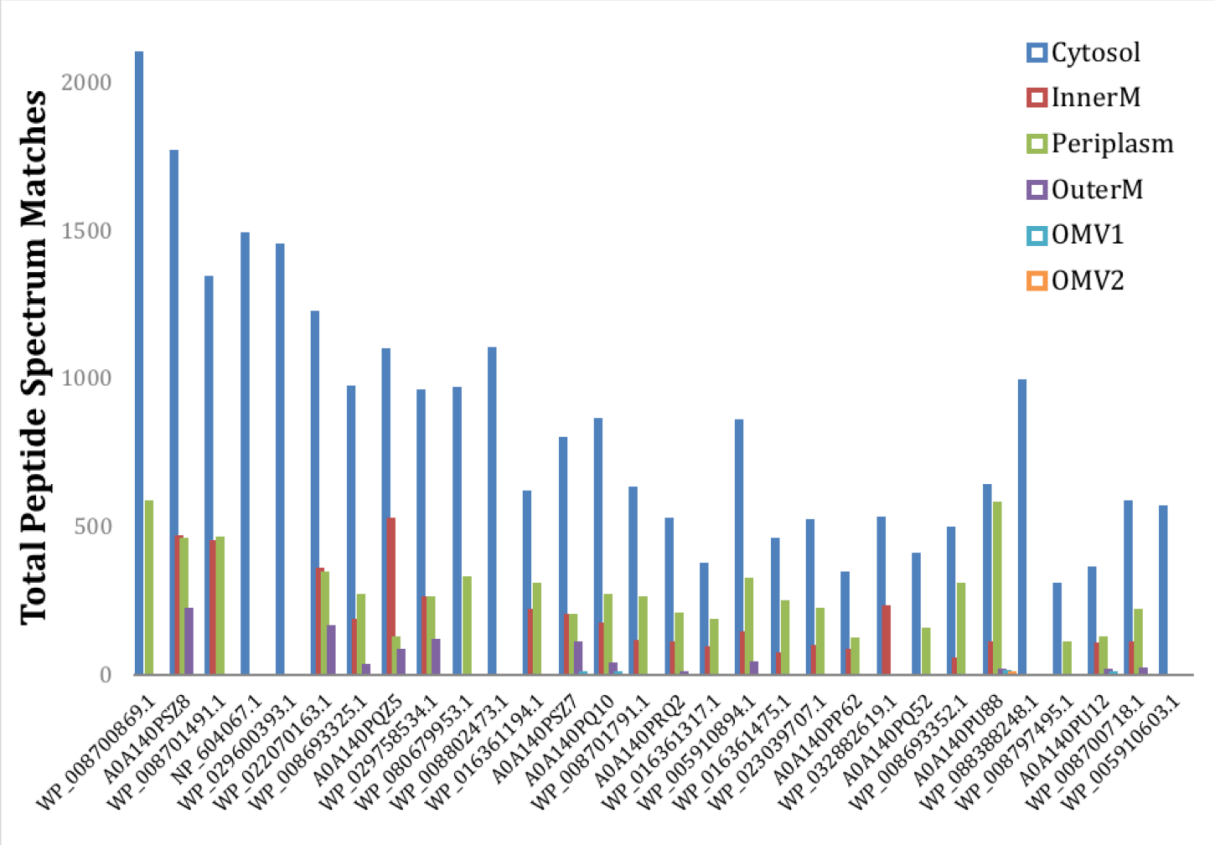


Figure 8 Purity of OMVs from *F. nucleatum* compare with cytosol fraction

The top 30 PSMs score proteins known in the cytoplasm fraction compare with the score obtain from all other fractions. Only protein A0A140PU88 was found in both OMVs. A0A140PSZ7, A0A140PQ10, and A0A140PU12 were identified in OMV1 only with small amount.

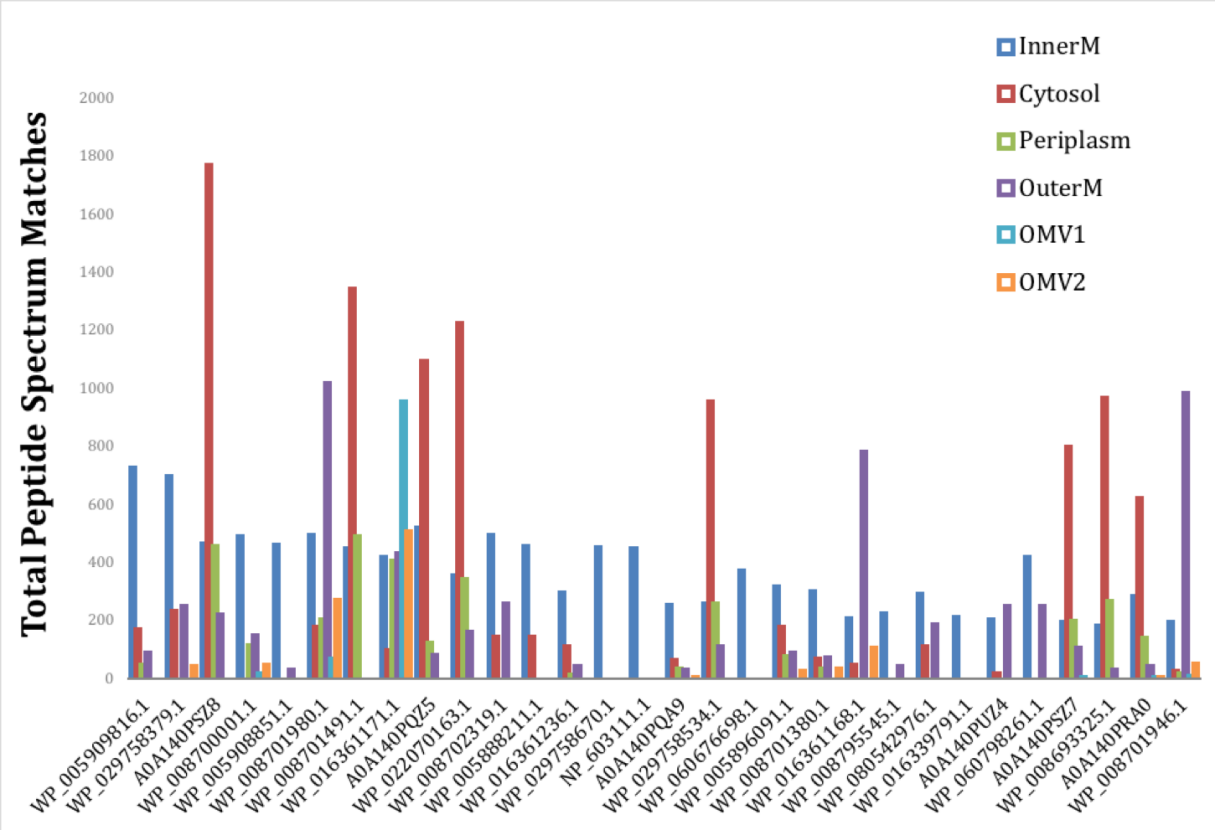


Figure 9 Purity of OMVs from *F. nucleatum* compare with periplasmic fraction

The top 30 proteins identified from the inner membrane fraction compare with score obtain from other fractions. Five proteins (WP_008900001.1, WP_008701980.1, WP_016361171.1, A0A140PRA0 and WP_008701946.1) was recognized in the OMV, periplasmic and outer membrane fractions.

Topology prediction of the proteins to OMVs

Among the 98 overlapping proteins of OMVs, 52 were commonly found in all subcellular fractions (Figure 10), 60 were predicted with signal peptides, including 31 with a type II signal peptide deemed likely to be lipoproteins (Table 2). Both Vaxign, which used the open-source database from PSORTb (v. 2.0), and Cello2go predicted the subcellular localization of identified proteins. Vaxign predicted the subcellular location for the 98 common proteins individually (Table 2, Figure 11). Many OMV related proteins were present in multiple subcellular fractions. Distribution was expressed as a percent of total abundance of a given protein (or a given class of proteins) found in a specific fraction. The calculation is as follow:

$$\text{Abundance \%} = \frac{\text{Individual Protein Abundance}}{\text{Total Protein Abundance}} \times 100\%$$

Although a variety of cytoplasmic proteins were predicted, outer membrane constituted the majority of the OMV in mass (Figure 12). Outer membrane proteins were the most abundant component of OMVs, which made up 72% of the total protein, whereas 17% consisted of proteins having unknown subcellular localization traits and 6% were determined to be periplasmic proteins. Cytoplasmic proteins represented only 5% of proteins enriched in the OMVs (Figure 12). This distribution pattern was virtually identical in separately run OMV1 and OMV2 topological analysis. Outer membrane proteins were the major ingredients of each OMV batch sample (not shown here), although the OMVs were found to contain a variety of protein localization classes. To further predict the localization sites of the OMV proteins to the vesicle membrane and lumen, relative abundance of individual proteins in the outer membrane and periplasm fractions (M/P)

was compared for the consistently identified OMV proteins with predicted signal peptide. Eighteen of the 60 proteins with predicted signal peptides were not present in the periplasmic fraction and only 9 of the proteins with signal peptide were abundantly (M/P ratio < 0.4) detected in the periplasm fraction. These findings indicated that the OMVs of *F. nucleatum* were almost solely derived from the outer membrane, although other proteins from subcellular fractions were present in small quantities in the OMVs.

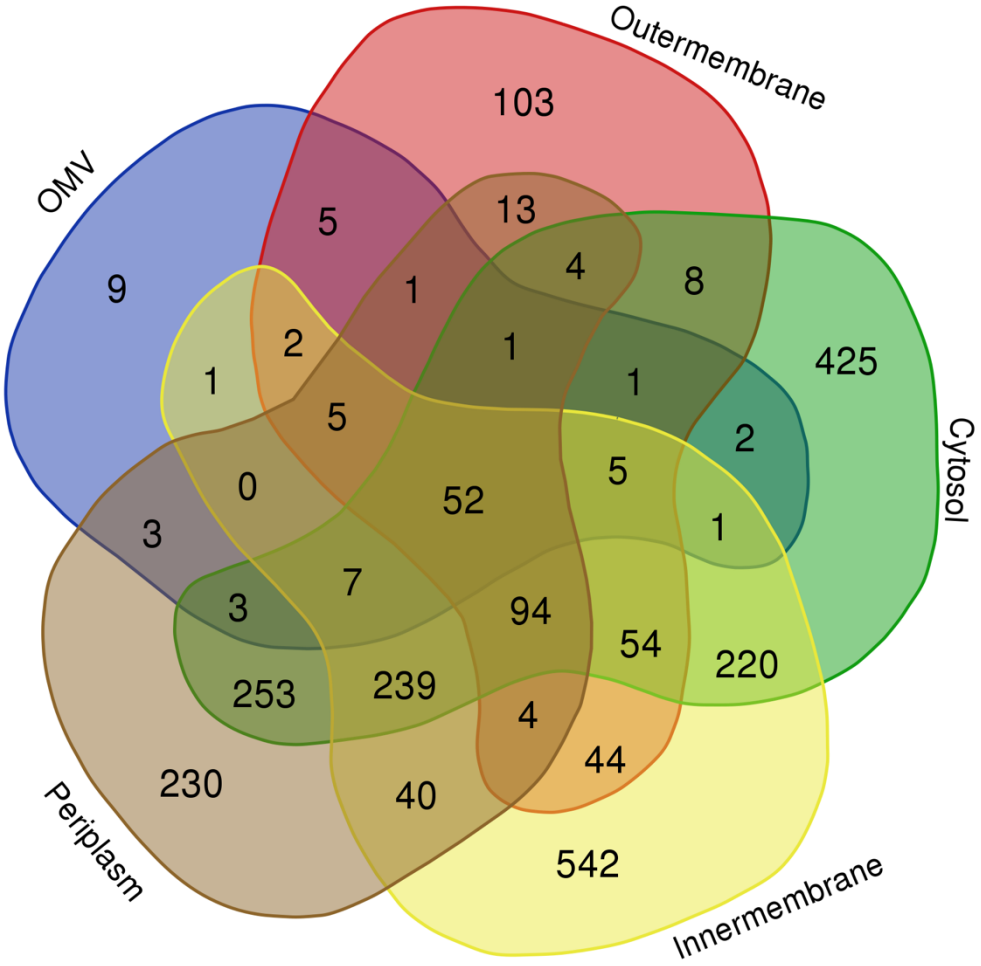


Figure 10 Venn diagram of proteins from all subcellular fractions

52 overlapping proteins are commonly identified in all subcellular fractions and OMVs. 72 proteins appeared both in OMV and outer membrane fraction, 73 proteins found both in OMV and inner membrane fraction, 72 overlapping proteins found in OMV and periplasmic fraction, and 69 are shown both in OMV and cytoplasmic fraction.

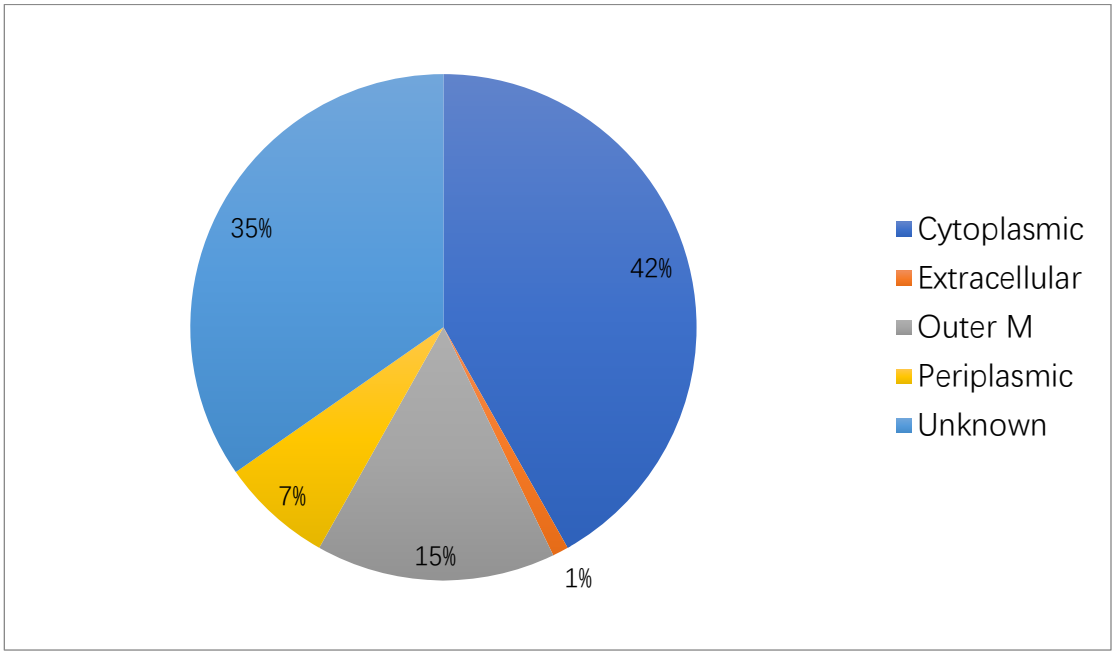


Figure 11 Topology analysis of OMV proteins based on proteins types

Subcellular location predication for 98 proteins on OMVs. 42% proteins are predicated as cytoplasmic protein, 35% are located in multiple fractions. 15% proteins are predicted as outer membrane proteins following with 7% predicted periplasmic proteins. Extracellular proteins only own 1%.

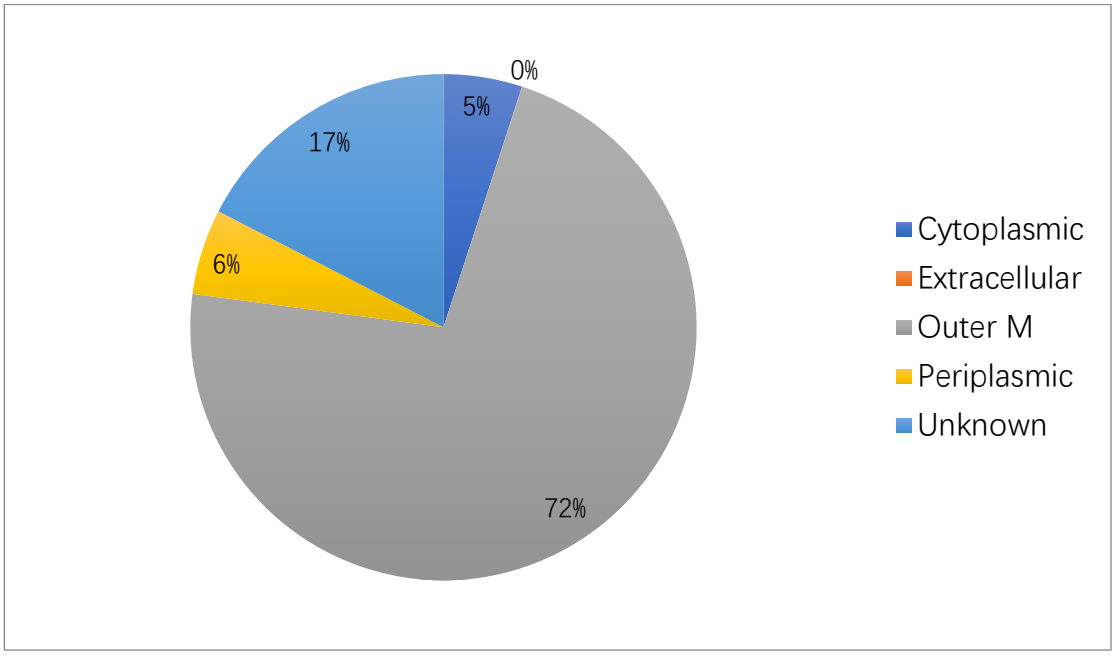


Figure 12 Topology analysis of OMV proteins based on protein abundance

Average mass abundance% from OMV1 and OMV2 for subcellular localization prediction. Outer membrane proteins were the most abundant component of OMVs, which took up 72%, following with 17% unknown location proteins and 6% periplasmic proteins. Cytoplasmic proteins only contained 5% richness in OMVs

Cargo sorting

By comparing the abundance of proteins from OMVs and outer membrane fractions, the relative enrichment of commonly detected proteins were estimated. FadA protein was found to be incorporated into the OMVs more than in the outer membrane fraction, while some of the autotransporter (AT) proteins were maintained in the outer membrane (Figure 13). Autotransporter proteins were highly present in the *F. nucleatum* OMVs and are also major composition of *F. nucleatum* outer membrane, which appear to occupy over half of the total mass of OMV. A single autotransporter protein (WP_016361171.1), with type II signal peptide and highest OMV/M ratio, was the premier component in both OMV replications, with an adhesion probability of 0.546 and predicted transmembrane β-barrel (Table 2). Although other autotransporter proteins appear at higher concentrations in the membrane fraction than in the OMVs, these proteins were predicted to display high adhesion probability for components of both cell fractions (>0.51) (Figure 13). Two YadA-like proteins were present in both OMVs and outer membrane fractions with mid to high adhesion probability (0.482 and 0.642). Numerous secretion proteins were predicted as highly likely adhesion proteins (square and triangle with lighter blue, Figure 13). The proportion of several identified non-secretion proteins were negligible; in contrast, some secretion proteins such as the Bacterial surface antigen protein (WP_023038391.1) and MORN protein (WP_008692599.1 and WP_008702680.1) were recognized only in OMVs (Figure 13). Overall, the secretion proteins were the major component of the *F. nucleatum* OMVs; the autotransporter proteins quantitatively constituted the majority of those incorporated into the OMVs and all the possible adhesion proteins were secretory (with predicted signal peptides).

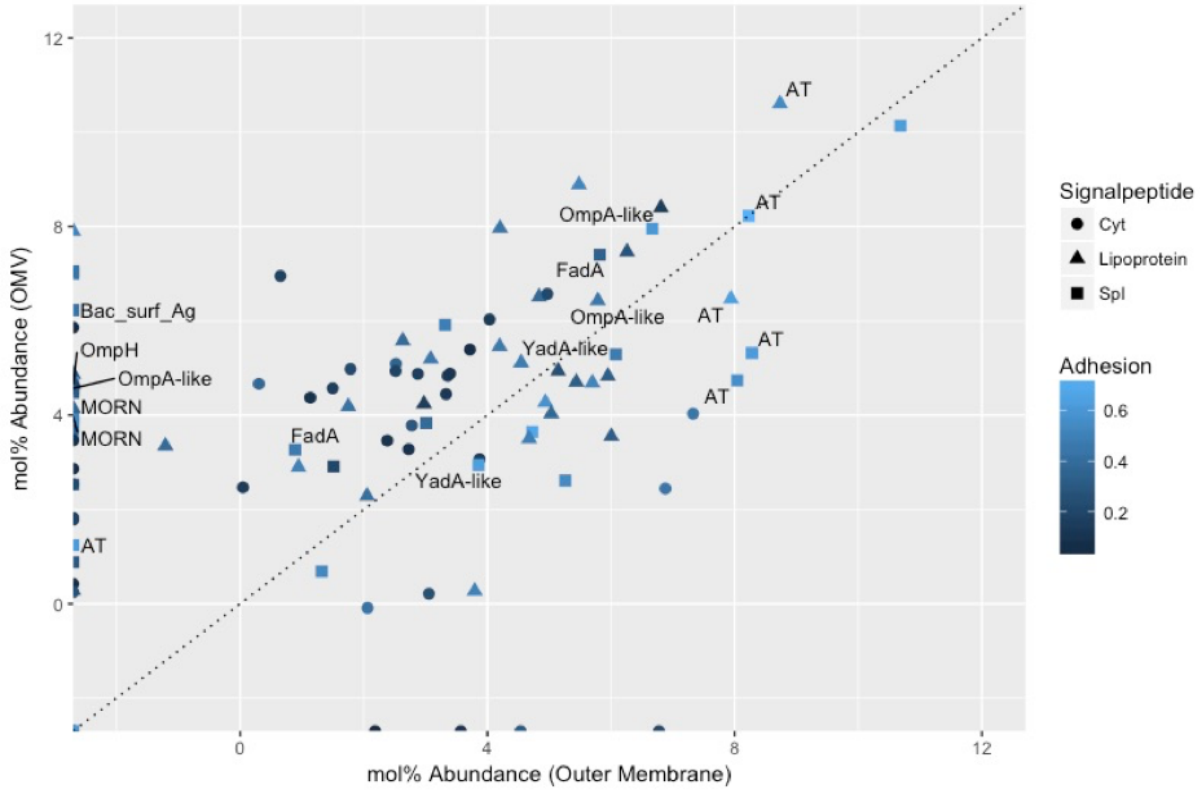


Figure 13 Cargo sorting to the OMV in *F. nucleatum*

The average protein mol% obtained from the OMV fractions compared to the mol% from outer membrane fraction. Proteins presented above the dashed line are relatively enriched in OMVs, oppositely, proteins below the trend line are deficient. The data for the figure can be found in Table 2.

Identification of Antigenic Proteins from OMVs

Reverse vaccinology was applied to investigate the potential for OMV-based vaccine development. In this approach, antigenicity of key targets such as exposed antigenic peptides, B-cell and T-cell epitopes was predicted using in silico analytical programs. The exposed antigen peptides for 16 predicted outer membrane and extracellular proteins contained in OMVs were forecasted by EMBOSS. A number of antigenic peptides (from 6 to 83) were predicted on each protein; an autotransporter protein (WP_008701556.1) was found to contain 83 anticipated antigenic sequences. This autotransporter was predicted as having multiple highly immunogenic epitopes within (Table 2, Table 3). B-cell epitope screening for selected outer membrane and extracellular proteins was achieved by BCPREDS. Non-overlapping epitopes were reported by both fixed (BCPred) and flexible (FBCPred) length prediction methods (Table 4). B-cell epitopes were successfully identified on all 16 proteins; the autotransporter proteins, WP_008701556.1 and WP_008701556.1 contained more antigenic epitopes than other fourteen proteins. T-cell epitopes of the top five most highly antigenic abundant proteins were predicted by Vaxitop, a vaccine epitope prediction system developed at the University of Michigan Medical School, to investigate the potential immunogenicity of *F. nucleatum* OMVs [46]. The p-value cut off was assigned as 0.05 for balanced binding probability and specificity. The T-cell epitopes against HLA-A2, HLA-A3 and HLA-DRB1*0401 alleles were predicted according to the previous study on tumor antigens recognized by T-cells [58-59]. The length of peptides against HLA-A2 ranged from 9 to 11, while for HLA-A3 the length was consistent at 9 residues. Among the five selected proteins, the autotransporter protein (WP_008701980.1) had the highest number of predicted T-cell epitopes followed by another autotransporter protein (WP_016361171.1) with 188 and 63 HLA-A2

peptides, respectively (Table 5). Likewise, the autotransporter proteins had the higher number of identified HLA-A3 and HLA-DRB1*0401 epitopes than other proteins.

Table 3 Antigenic Sequence Prediction

Accession	Hit Count	Sequence	Residues	Score
A0A140PNL7	7	IIASCMLALSLVGCTGF	6->22	1.178
		ANAVAQYLIAQGVSSSRI	161->178	1.156
		LIGAAVGSLVG	56->66	1.147
		NALVINLPGGVTFAS	96->110	1.134
		GAAVGALAGQI	37->47	1.109
		SGFYSALNGIAQSLN	117->131	1.066
		QGTQAQVK	85->92	1.039
A0A140PP98	6	KLYAVLLLALLVTACSS	4->20	1.242
		LEDIIVF	49->55	1.106
		NLELSVKRA	124->132	1.098
		LILSMPELILFDFDKYIVKDK	69->89	1.082
		KPSLSTLAKA	91->100	1.080
		KDFLISRG	136->143	1.072
A0A140PUD9	10	LALVLGSLLVVGSVASAKEVMPAPTP	4->29	1.219
		PEKVVEYVEKPVIVYRDR	31->48	1.164
		SYELYMLPTFQVSYKPTDFVKLYAAAG	316->342	1.135
		YTWHQYKVIAN	298->308	1.116
		YYGALEAYLYQHTPL	265->279	1.114
		AEASVLFDFADYL	164->176	1.108
		SVDVQYR	59->65	1.099
		TLDVRVR	106->112	1.097
		EYTLPLGFS	230->238	1.086
		SSKVKAT	142->148	1.060
WP_008691343.1	19	KLSILIVAGILVGCA	4->18	1.180
		NIAVVAKY	157->164	1.162
		VEEVPVDEGKVISK	48->61	1.152
		AIILRVAKL	243->251	1.134
		SFRSVYQGQYILSTY	133->147	1.131
		EVNYLFLKYA	205->214	1.130
		KVEVTKYLRLLKQ	392->404	1.124
		KSEVLTK	414->420	1.115
		KEVYDVIMQ	169->177	1.112
		EYLSIVKGN	70->78	1.109
		KSALQEYVGHK	271->281	1.105
		NEVVRDKIKAVIDGT	304->318	1.090
		MHEALAGV	223->230	1.084
		TVGLGETLIFPL	95->106	1.083
		ENAVTLY	187->193	1.075
		DIFVPSE	255->261	1.071
		EIYYNIASSYAK	377->388	1.070
		IKLTAS	112->117	1.054

		SIKKVEDS	82->89	1.051
WP_008700150.1	14	ILFIILCLFLISCSNL	4->19	1.226
		VGTIHYFLCMRG	388->399	1.154
		NVELFQIYIAS	87->97	1.129
		VQNVELTFK	32->40	1.117
		AYALEIIDYVI	144->154	1.116
		YITLAEV	193->199	1.110
		KPIITVFNF	370->378	1.097
		VRNTVKNVK	412->420	1.088
		SIGYSVSNGIELFIE	324->338	1.087
		IELYSS	175->180	1.076
		YLRSAQNLYYKADKIYSRYQ	208->227	1.063
		QKYVSNY	281->287	1.063
		IFEYYKK	67->73	1.053
		QRKVAEK	445->451	1.052
WP_008700266.1	19	SVSLKLIIFSFLLVSSSIAYSAAPTI	4->29	1.176
		KQSVAVGLSYY	612->622	1.161
		ESSVVLGHGSSVSESEVVSVGT	307->328	1.149
		PQYVVQNEVKRLTV	665->678	1.136
		KQLYQVAKA	354->362	1.124
		GFTVKLG	648->654	1.117
		EVNHKHLLIN	258->267	1.114
		KDEVNHVGSLSAALAGLHPMQ	573->593	1.112
		YSSAVGYE	93->100	1.095
		QNSVALG	497->503	1.093
		DIDVAAWKAKLGVGSGGVDLTAY	366->388	1.092
		SNEVSVGSA	511->519	1.091
		VSTVDSSAF	117->125	1.082
		PTQVMAALGHY	599->609	1.081
		KHSSAVG	77->83	1.077
		IVNVREA	335->341	1.069
		DGDVSAT	530->536	1.064
		DPNVVTGN	460->467	1.062
		SAGVAIG	630->636	1.060
WP_008700372.1	13	SVSLKLIIFSFLLVAGSVAYSAPDF	4->28	1.176
		KQSVAVGLSYY	419->429	1.161
		YNEVPQYAIQDEVKRL	468->483	1.122
		GFTVKLG	455->461	1.117
		EVNHVGSLSAALAGLHPMQ	382->400	1.112
		KVFDYD	261->266	1.097
		SNEVSVGSA	318->326	1.091
		VGTVNTV	221->227	1.088
		SSTVGLYNKA	190->199	1.086
		ASGVYSSAF	73->81	1.085
		PTQVMAALGH	406->415	1.081
		KAKVDSQ	498->504	1.062
		SAGVAIG	437->443	1.060
WP_008701093.1	11	KTILFSLILIFTFVSCSTLHTVVSKRNLDVQ	6->36	1.165
		AKYWLQANILKVDKVNL	91->107	1.128
		TVFVQIR	54->60	1.115
		AGAAIGTLADALVDDT	143->158	1.112
		LGKVIA	231->236	1.092
		AIPLLED	223->229	1.092
		YAMVTDILIT	160->169	1.089
		STRVLSTA	206->213	1.086

		ITNVLTSK	73->80	1.074
		IGGVLGA	122->128	1.067
		DAALGA	115->120	1.049
WP_008701556.1	83	KGLVHDVLVSYKKSGYHFIQD	502->522	1.231
		EKVLVVKTF	232->240	1.190
		AYGVVYLH	1867->1874	1.188
		NGKVVPPTK	1605->1613	1.159
		QIKLLPAVGSIGT	1530->1542	1.156
		RKYLVVNEI	1944->1952	1.155
		AKGVYLVGN	1300->1308	1.149
		LTTLVGVA	2036->2043	1.146
		RGGVGLR	2110->2116	1.144
		YTPVLGASIL	246->255	1.141
		TSRVITVPQA	261->270	1.139
		NYIPYSLIESNLK	994->1006	1.136
		IIEVEVSAGVKPR	189->201	1.136
		RVLLLANGT	956->964	1.135
		SIKYSLGMAILFLMLGVSAFS	21->41	1.132
		PILLAQA	1794->1800	1.127
		YKNYSSIISYVMG	340->352	1.126
		NSLVRLERV	648->656	1.123
		IVGLYGYQ	1255->1262	1.122
		THGVNFTSSIVNLLSSN	316->332	1.119
		IGAVGAV	1512->1518	1.113
		SDGVTYIRVK	819->828	1.113
		NLNLLSPKTKVDLIFGI	1669->1685	1.110
		SYKYYFGT	2026->2033	1.109
		TNAAVPNPDYLSVSPG	1616->1631	1.108
		YANVQQRVQ	1809->1817	1.107
		IRPYGALKVE	1982->1991	1.104
		AFYYKL	918->923	1.103
		GLIQLLGEH	1191->1199	1.100
		DQVVKSP	97->103	1.097
		GARLELIQLT	84->93	1.097
		QIYSASLTWI	1719->1728	1.095
		NIYVGAS	1416->1422	1.094
		KNLHPVIKN	460->468	1.094
		SVGLFSSQ	832->839	1.091
		QSSIVNEK	1027->1034	1.091
		NQVVLG	1343->1348	1.089
		GMYLDQH	1490->1496	1.088
		IAPIQNV	275->281	1.087
		YGIYSSGSVINY	1377->1388	1.087
		IGMYVDTS	1652->1659	1.084
		VDIVSE	638->643	1.083
		NVAYKFSKL	1312->1320	1.082
		NAKVIQIGEKILKPY	1692->1706	1.081
		RIPPLPDF	220->227	1.080
		PLKLLGK	551->557	1.078
		EQVAIA	1042->1047	1.078
		RGSVENLQIKLQKS	61->74	1.077
		FGLVNI	1327->1332	1.077
		QAKIGLFKSK	1910->1919	1.076
		GDIFVGYN	1933->1940	1.076
		IKLIGAR	1481->1487	1.075

		SGALSSLAGKL	977->987	1.075
		KITVGDSGIGVYSSK	1270->1284	1.075
		IGIYYS	1129->1134	1.075
		SKYYTYRIG	1957->1965	1.073
		RIVLGNA	1218->1224	1.073
		SQLITAH	680->686	1.072
		KVKLFRT	794->800	1.072
		KSIFRIY	604->610	1.072
		SNAKVIAEE	851->859	1.067
		NQVFQKL	1780->1786	1.066
		RYGVEAL	1769->1775	1.065
		GSLIHF	411->416	1.064
		DFLFYG	770->775	1.063
		NDYLSIKP	2012->2019	1.062
		AGIVDYK	1858->1864	1.061
		AGDLKIG	1290->1296	1.060
		DKVSLT	1057->1062	1.059
		YELIAT	583->588	1.057
		DLQLAEE	181->187	1.055
		LKNILSS	1157->1163	1.055
		INAVANK	862->868	1.054
		GMVFYN	561->566	1.053
		AKIPYTN	1744->1750	1.050
		IEVIGKKS	432->439	1.049
		KISVGKK	1091->1097	1.042
		DRVVKN	1408->1413	1.040
		NNVLFT	877->882	1.038
		PTVKID	1589->1594	1.035
		KTVVDE	1580->1585	1.033
		RSIAKRF	13->19	1.022
		KEFKYL	1825->1830	1.020
WP_008702403.1	46	IRKVVVSQV	359->367	1.213
		MEKLIYYKLV	898->907	1.184
		LFRVILLDSQ	406->415	1.178
		YLIISLLASYSIVFAG	4->19	1.175
		VAKIYFLE	63->70	1.147
		KAALYLIEI	236->244	1.140
		STKVYLNRV	539->547	1.128
		YNQSIYHIFAINY	332->344	1.125
		LGIVYYNLK	168->176	1.122
		LFTKLYAVK	880->888	1.113
		YLSLAQDE	496->503	1.110
		VSYLRASAIYR	200->210	1.108
		KVRFYLEK	253->260	1.107
		AASLYKEVY	633->641	1.107
		FQKVLASGDQ	527->536	1.106
		YGRLFAVSP	393->401	1.105
		YYSAQYYAAK	759->768	1.105
		KTFFAVAQDFL	128->138	1.105
		EYILYRLG	823->830	1.103
		IEKYIAELP	744->752	1.099
		KDVYLYTGD	441->449	1.097
		IITLKLQKK	731->739	1.094
		VISSLMSTLLEQK	476->488	1.092
		EQAYYKYIMTLS	651->662	1.092

		MALTDYKAVYSK	841->852	1.092
		PKAVKEYETLLQYDKYKAY	771->789	1.090
		GYEVYGKF	611->618	1.087
		NYFLAE	549->554	1.085
		GDLYVTK	275->281	1.085
		YASYLARANA	96->105	1.084
		QIYKTYLNQ	462->470	1.083
		SVNYLKGI	505->512	1.075
		KLIYGEAYSLYKK	301->313	1.073
		KIGLSYFRLG	585->594	1.073
		ARSYYSKI	599->606	1.070
		EAVQAGEQYLSKI	558->570	1.068
		KKFQSL	321->326	1.068
		IRIIANSAYQVGNY	373->386	1.068
		LKAAEIY	862->868	1.067
		KAIYWLSEF	180->188	1.066
		AKKYYGTYAT	807->816	1.066
		KFLQKY	45->50	1.063
		MKVYPN	678->683	1.060
		MFYLRE	114->119	1.044
		IKYFEE	218->223	1.041
		KKLYDEL	915->921	1.034
WP_016361171.1	41	IALASILFVSCG	13->24	1.187
		GSGVKVAILDANFVNAVRS	103->121	1.148
		GEEVLEVVRDLEY	184->196	1.141
		KMLSAYVELG	941->950	1.139
		KLQAYVTQA	1007->1015	1.138
		MVGVSLYGKQDLPYGFY	890->906	1.134
		NVEVKDGKVVATL	706->718	1.133
		SVKVLGKI	654->661	1.129
		LPLPFAKVHSG	288->298	1.127
		EVWFSVLGSA	825->834	1.127
		VAAILPTQEVYEA	238->250	1.127
		SGTLVLNPKAIVGYDVDNF	584->602	1.126
		EITPVFGVYGNVDFR	1057->1071	1.124
		GSKVIGGTLEIHQVHASPITVG	561->582	1.123
		GSSYAAPKVTRAAALVYDK	428->446	1.123
		NGTVSVLSNGY	665->675	1.120
		TLGVALNYSYAK	865->876	1.115
		LKASLYTA	79->86	1.104
		GGDYVAYA	636->643	1.103
		LTDVYTDVEII	149->159	1.101
		ADSVFSAGLQ	1078->1087	1.098
		LMGATVQS	764->771	1.098
		AGRLGLSHISSK	908->919	1.094
		WISVIGA	348->354	1.093
		SVEAGLPYFD	332->341	1.092
		IYASAQALTFS	788->798	1.090
		HPLVPRR	139->145	1.088
		SNFLVGARAE	991->1000	1.086
		LGTLHL	547->552	1.078
		TPSVPKPS	43->50	1.076
		GSSVKQKFYGVK	1031->1042	1.075
		VDKLSSA	382->388	1.074
		KTILGSVGWDYTYV	216->229	1.073

		LMIPYFRD	303->310	1.072
		YWSIAADAYGVD	394->405	1.068
		KATLPAGKT	521->529	1.066
		FTPFIGYSQDY	956->966	1.053
		NQSVKIFNQ	257->265	1.051
		FPTVTNP	62->68	1.051
		AENVEKVFED	739->748	1.051
		SSGIKVK	615->621	1.042
WP_016361319.1	17	LLLLLAILSSGVYASSIDHIQTYSPDYLSNQ	4->34	1.209
		RRDLPAILSVGASYK	328->342	1.168
		VSGVDVLTDLAQFKPLGIY	121->139	1.164
		TLAITASVAHFIY	437->449	1.138
		TKFQQRVSGLVKDGLVGDI	238->256	1.134
		SFSIAGRVVHGSR	167->179	1.119
		GLTTFYPQYIIDSK	313->326	1.114
		YFHVGIQ	60->66	1.112
		ALNSVTLGAGVRYN	420->433	1.105
		QAKAQQITQEVSKAVDAS	197->214	1.100
		KAILNQLIPNIT	82->93	1.094
		DNYLVSGT	345->352	1.093
		GLSAAQIAAIK	219->229	1.089
		IDSVSPFYN	40->48	1.084
		EHQLQTTDVLGQT	299->311	1.083
		LGRAFSVNDK	156->165	1.078
		GFQLGINYK	266->274	1.075
WP_017139894.1	40	KDEVTHVGSLSAALAGLHPMQYDPKAPTQVMAALGHYKNK	1515->1566	1.160
		QSVAVGVGYYFN		1.136
		TPQYVVQNEVKKLSVENK	1606->1623	1.132
		NFKLLSLGIFIILMQNSLVAAD	3	1.112
		FKSGVSNSVAIGNGSVASTSNEVSIGSDTVKRRLTNV	5->26 1446->1483	1.106 1.098
		KSKVEVQDE	2	1.094
		NVGFTLKLG	1626->1634	1.093
		KNLVAGLA	4	1.087
		NAGVAVGSE	1588->1596	1.087
		NKNLVAGLS	6	1.086
		VGSKNLVAGA	295->302	1.083
		SNVVGGA	1572->1580	1.080
		SAKSTDAVNGKQLYNAVKGVE	0	1.075
		KNIVAGVG	238->246	1.072
		KNLVAGL	278->287	1.068
		KGSSSTVIGYKNTV	85->91	1.059
		NKVLGLNNT	1488->1505	1.056
		NNVLGLDNE	8	1.049
		NLALGYKNI	771->778	1.049
		NYIVGTF	267->273	1.049
		ESTVVGN	1377->1394	1.042
		KNIVAGA	0	1.042
		IGNKASVYG	212->220	1.039
		NLILGA	226->234	1.039
		IAGVGNI	975->983	1.037
		IKNLVNG	44->50	1.034
		NKVAGAENLVNG	1226->1237	1.033
		SSALGIGNT	2	1.032
		RNLVNG	799->805	1.031

		DNVIAG	1083->109	1.029
		NKIVGSF	1	1.028
		GKYSTA	254->259	1.028
		GSGVEYN	760->766	1.027
		TDNYVAGR	854->860	1.022
		KNLVAG	863->874	1.020
		KNLVKGS	1367->137	1.016
		VAGIGNKV	5	1.015
		KNIIAG	841->846	1.015
		SNKVAGA	155->160	0.987
		SNKVAGA	72->78	
		KGISNT	1238->124	
			3	
			1598->160	
			4	
			1001->100	
			8	
			309->314	
			813->819	
			956->963	
			785->790	
			575->581	
			351->357	
			201->206	
WP_023038391.1	28	DGGVRIVVDVVEKENVVDLL	82->101	1.203
		ILIAFLFVISLTSFSTMVNLPIKSIEVVNNQQVPAS	4->44	1.183
		LIKNT IIGFVVFAD	621->629	1.161
		KSIVISSVKFI	117->127	1.158
		WLWSVYPYISY	505->515	1.156
		IDLILLID	345->352	1.153
		ALSFEVVENQ	182->191	1.148
		FEDVVLQPT	71->79	1.145
		LHISIVEGIVRKIEVKKM	258->275	1.138
		VKNIVITG	193->200	1.133
		AYKVVGGI	568->575	1.132
		PTSGVYGKFQVEGG	524->537	1.120
		ILGLYQAQGYTLVNI	233->247	1.109
		FSEVKPDAQVA	167->177	1.101
		SQKLVATI	605->612	1.100
		TSELLDITQL	134->143	1.099
		EKGVAVNT	103->110	1.091
		QGGVAYG	360->366	1.089
		ELFHEIDTVGFKVN	435->448	1.084
		FKNVKYEARSVPG	328->340	1.065
		LKTKDYVIDR	291->300	1.063
		TEALLADFN	55->63	1.062
		NVTLELRTYHRG	549->560	1.061
		GFALDFYDP	402->410	1.057
		RRLLATG	159->165	1.056
		PIGPLRF	667->673	1.053
		IFNVKEYDATVD	309->320	1.049
		GWPVGH	676->681	1.041
WP_060675827.1	7	MVMLFLLVFSLPALAVQALT	13->32	1.205
		EFKLVQREA	170->178	1.129
		EAVKAKLLEFG	126->136	1.124
		DKSIVKAQYFDLLKN	76->90	1.114
		MTIVLDE	62->68	1.069
		ITIVGHT	102->108	1.069

		DRIVGIE	140->146	1.063
WP_060676683.1	12	LALVLGSLLVVGSVASAKEVMPAPTP	4->29	1.219
		PEKVVEYVEKPVIVYRDR	31->48	1.164
		ASYSVYMLPTFQVAYKPTDFVKLYAAAG	317->344	1.135
		YYGALEAYLYQHTPL	266->280	1.114
		AEASVLFDFADYL	165->177	1.108
		SVDVQYR	59->65	1.099
		TLDVRVRNYHSL	106->117	1.097
		YTWHQYKVI	299->307	1.097
		SSKVKAISRLEYK	142->154	1.095
		EYTLPLGFS	231->239	1.086
		AVELSFD	285->291	1.078
		YAPYDG	210->215	1.050

Table 3. Predicted exposed antigen of OMV carried outer membrane and extracellular proteins

The exposed antigen peptides for 16 predicted outer membrane and extracellular proteins carried by OMVs were forecasted by EMBOSS. 6 to 83 antigen sequences are predicted on each protein; specifically, autotransporter protein (WP_008701556.1) contained 83 anticipated antigen sequence, which is most among the 16 proteins.

Table 4 B-cell Epitopes

Accession	BCPred Epitope	BCPred Score	FBCPred Epitope	FBCPred Score
A0A140PNL7	TGGAAGGAAVGA	1	TTGGAAGGAAVG	1
	RDNQERELKAKL	0.996	RDNQERELKAKL	1
	VGCTGFEAGNGG	0.978	VNGYTDNTGKDA	0.995
	ASNSTPDGRLQN	0.975	PGGVTFASDSAN	0.983
	IIGKDTKGTLIG	0.953	SNPIASNSTPDG	0.971
			VGCTGFEAGNGG	0.971
			NQELSQRRANAV	0.941
A0A140PP98	TEANKKPLEDII	0.997	EDTEANKKPLED	0.999
	CSSTKTNRKTNV	0.961	GEQNPADTNQTA	0.999
	IEGYGEQNPADT	0.951	TKTNRKTNVGVD	0.995
			FNEEGVTIRREG	0.97
A0A140PUD9	KEVMPAPTPAPE	1	KEVMPAPTPAPE	1
	RGTKGDGASDEY	0.996	KSEDTDKDWANS	1
	ENKTKSEDTDKD	0.991	VVEYVEKPVIVY	0.999
	SKGNKKGQYYGA	0.979	NWSGHSNGGEGR	0.999
	TNNSEAKNWRWQ	0.971	TFDGYNGGTNNK	0.999
	VTFDGYNGGTNN	0.971	PAWRPNGSVDVQ	0.998
	SRLEYKQKSNDG	0.97	RGTKGDGASDEY	0.995
	DREVAPAWRPNG	0.967	NGSGYERASYEL	0.993
	FEGGYDPYTWHQ	0.948	RNFAVTNNSEAK	0.99
	NQTLDVRVRNYQ	0.946	SNDGAKHAEASV	0.989
	VYNGYDRYQGSF	0.941	LSFDFEGGYDPY	0.986
	KFGLRPGYAHNW	0.914	TTANVNFTKNQT	0.97
			LNFESEYTLPLG	0.946
			NVYNGYDRYQGS	0.926
WP_008691343.1	EIGGEPQIPGET	1	ETKIVEEVPVDE	1
	TASPKNTNTKIN	0.992	KLTASPKNTNTK	1
	NPELKSALQEYV	0.989	EKAKKEAPETAN	1
	FLKYAYDTRNSS	0.961	DSANESKDLENA	0.999
	EKAKKEAPETAN	0.942	KKVEDSVENDYT	0.999
	PNSEWAKKSEVL	0.929	SETYKTNNPELK	0.999
			KAEPSGESYYDK	0.999

			KYAYDTRNSSLM	0.998
			NLGGKRDSGPIR	0.992
			EFPNSEWAKKSE	0.988
			AKYDFSEKEVYD	0.978
			EKMQDKKPEIYY	0.952
WP_008700150.1	FKKLKEKKKIEI	1	EVEEKQESNQYL	1
	NTSSNTNTEQMS	0.999	NFKKLKEKKKIE	1
	NSNNSEISSRIN	0.996	NTSSNTNTEQMS	1
	FTKTTTTIEEVG	0.994	EYRNAGSSKAKY	0.999
	GSSEIQRKVAEK	0.969	TVFNFTKTTTTI	0.998
	NNINNDVTVRNT	0.919	SLNSNNSEISSR	0.998
	IFEYYKKLKNSE	0.915	RDSDNRTLGSSE	0.995
	PPNSNYRDSDNR	0.911	NDVTVRNTVKNV	0.987
			ASDNSEYSREFQ	0.984
			VAESRPDLDKSK	0.965
			NYYFREAQKYVS	0.944
			QEVNSNINSMVS	0.931
			DIFNLALKTNKE	0.917
			CSNLYKANKAYE	0.906
WP_008700266.1	LGVGAGGGGGLV	1	GVGAGGGGGLVN	1
	GNDNTINGNNNF	1	GSSVSESEVVSV	1
	GTPDGGRRIVNV	1	NQGKENQELKEE	1
	GAFGDPNVVTGN	0.999	LGNNSTVSSSNE	1
	NQGKENQELKEE	0.999	NATSNDATTGKQ	1
	TATGTGSTGLGI	0.998	GTGSTGLGIDNT	1
	GGGNSYMIGNGN	0.996	INNDASGEDSSA	1
	GAFGIGTYDEVN	0.989	FGNDNTINGNNN	1
	GNNVTIGSGIQN	0.98	GNTANEENSSAF	1
	KNEVSTVDSSAF	0.976	KIDDVKDEVNHV	1
	WKAKLGVGSGGV	0.972	KSGAFGDPNVVT	1
	PTIEAGTGTSST	0.957	GNTASGEDSSAF	1
	SNDATTGKQLYQ	0.936	VLGNNVTIGSGI	0.999
	KIDDVKDEVNHV	0.921	GNQNVTSGEGSS	0.999
	TAIGYKNNVSGN	0.912	GHKNEVSTVDSS	0.999
	AGVAIGGERRVK	0.91	NTKNKEQDEKIK	0.999
			GVEYNETPQYVV	0.998
			NGGGNSYMIGNG	0.997

			GAFGIGTYDEVN	0.994
			DGDVSATSTDAV	0.991
			SGQSSSSFGHSN	0.991
			GRDSNGRKIESS	0.989
			GGRRIVNVREAR	0.989
			GERAYSSAVGYE	0.971
			QYDPKAPTQVMA	0.97
			GFNNEATSLEST	0.963
			PTIEAGTGTSST	0.942
			DLTAYSKRDASN	0.927
WP_008700372.1	AGTGTNSTVAGV	0.998	KQAKENQELKAK	0.999
	GAFGVGKATWNK	0.998	GNGSTVSSSNEV	0.997
	NNKQAKENQELK	0.993	LGNNVSIGSGIN	0.995
	KIDNVKDEVNHV	0.948	GVNNEATKSYSS	0.993
	LSNTATGMNSSA	0.924	ADGEVSATSTDA	0.987
			RNSSAFGYNNAV	0.986
			GTVNTVSGKNSS	0.984
			SSAFGYNNAVTG	0.979
			SNTATGMNSSAV	0.976
			DNVKDEVNHVGS	0.975
			SGTDSSAFGDQN	0.975
			QYDPKAPTQVMA	0.97
			GSGVTYNEVPQY	0.933
			KAMQNSSSTGIE	0.908
WP_008701093.1	GAGIGGVLGAQR	1	QGNSGSMTSTST	1
	QGNSGSMTSTST	0.993	NLDSDNGFSDAA	0.999
	VNNPAEAKYWLQ	0.962	RIVNNPAEAKYW	0.995
	KTGRNVQTSTRN	0.927	FNSAIPLLEDEL	0.988
			GAGIGGVLGAQR	0.986
			VQIRNTSGKNLN	0.983
			GRNVQTSTRNSV	0.975
			IWLEPAAANERT	0.913
WP_008701556.1	RGEFEGDGGYKG	1	DTNTNTLVNTGN	1
	GGYYDTNTNTLV	0.999	KSKAENEEQLKG	1
	GAPYKGRGDKPN	0.997	TKVDTNAAVPNP	1

LGKGIGWYAGIA	0.996	EEPIIEVEVSAG	1
EYYGDDPDDSTK	0.994	NSDNVTGITIQN	1
LTGDKSTGIYGK	0.994	RGEFEGDGGYKG	1
AEEPIIEVEVSA	0.991	NSVQTREEITTS	1
TGSNSNTVGGAY	0.987	EYYGDDPDDSTK	1
GIEAAEYTNAKV	0.987	NSDNMAATNNKK	1
TNAAVPNPDYLS	0.986	KNSGSITGGKNI	1
GYQITTTNTSKI	0.985	LYGYQITTTNTS	0.999
ITVGGTNYSNNR	0.98	LLEDNEATTMGG	0.998
GHNVRGGVGLRV	0.979	GAPYKGRGDKPN	0.998
ETGQIIGGGHLP	0.978	NAKNNSLVRLER	0.998
YVGASQPREHYS	0.969	LTGDKSTGIYGK	0.997
ENGNTKEADKVS	0.968	KFKDIGNSKEEQ	0.997
AMYLLEDNEATT	0.965	TKGDGVANGYDV	0.997
KIDTKKTPTEVE	0.963	KPYDDNNSLNWT	0.997
GSITGGKNIVGL	0.957	LLNDNTGKISVG	0.997
EKDGTAFYYKLP	0.953	TGSNSNTVGGAY	0.996
NSDNVTGITIQN	0.946	NRKNNGKIGSIG	0.994
LNDNTGKISVGK	0.946	NHTTLRSNGDQN	0.993
NRKNNGKIGSIG	0.943	YVGASQPREHYS	0.993
NSVQTREEITTS	0.942	NGNTKEADKVSL	0.992
LDSNTDSYNKMQ	0.938	GGGHLPTTEEDY	0.992
KSGIINFKDNNV	0.936	NQRIGVNANVGY	0.991
GIYYSDKDKNGN	0.931	YKYYFGTKSLTT	0.991
FDEMKGKQYANV	0.927	YTNPIQNLNLLS	0.989
KFKFKDIGNSKE	0.909	GGTNYSNNRSAI	0.989
GERGIGMYATGY	0.903	IPPLPDFEAPNE	0.987
		EITGSQNYIPYS	0.987
		YGMNAKYGGKII	0.987
		DVNLDSNTDSYN	0.986
		SAAINFSYPWKG	0.983
		GKGIGWYAGIAH	0.982
		SAPNKITTLTIG	0.982
		TVPQANDKSIAP	0.981
		GSIGIYYSDKDK	0.98

			INAVANKSGIIN	0.979
			TAFYYKLPNNGN	0.976
			TFSGDATYTPVL	0.972
			THGVNFTSSIVN	0.969
			VDYKNTAYGVVY	0.969
			DSITEDNATSGG	0.968
			WYNLRGEKEDRR	0.967
			NTPKNLYTEGTI	0.959
			NIQGDLKAEIEK	0.958
			NGGDSSHFGTTD	0.95
			GNIEVHGERGIG	0.95
			NEIFNAKSKYYT	0.948
			VTYIRVKDKKSV	0.947
			PYTNFAKKSDSY	0.938
			LKKEWRTASKDS	0.937
			NVNISNGKIDIT	0.93
			QQYGFFSNNNGG	0.929
			GIEAAEYTNAKV	0.915
WP_008702403.1	KNNTKEAKKLYD	1	NNSSANTNDASV	1
	AKENNEAAKKAY	1	EQKKYDEMEQYL	1
	AELPIEEERTYY	0.999	YFDKKDLSKAKK	1
	YKKGKYEEALKK	0.998	AKKDPKAVKEYE	1
	GDPERAEQIYKT	0.997	ELPIEEERTYYS	1
	NNSSANTNDASV	0.953	YPNSNLRGTITN	1
	TKENYEKALDYY	0.946	GDPERAEQIYKT	0.999
	AVKNNKELKAYS	0.945	EEKQSQKEEYAS	0.999
	LSLAQDESSVNY	0.944	AKKAYTEIINNK	0.999
	EKQSQKEEYASY	0.936	YDKQENVAEAEK	0.997
	TPASEYSRKAAL	0.925	LVKNNTKEAKKL	0.997
	SNRKDEAKVRFY	0.912	NMTPASEYSRKA	0.996
	NEKNYEKAASLY	0.9	KKGKYEEALKKF	0.993
			LLNNSNSAMFYL	0.987
			LAQDESSVNYLK	0.987
			KNNKELKAYSME	0.984
			YFLEKNYKKAEE	0.982

			VYYNLKDYDKAI	0.977
			EFSKAMPNENKE	0.967
			RNFKTALTESEK	0.96
			YVTKENYEKALD	0.957
			TVNKNPDEYILY	0.952
			IESGDTDKAIES	0.947
			AKVRFYLEKLKG	0.942
			LGKYDEAEADFQ	0.934
			LLQYDKYKAYAS	0.918
			IMTLSLTGNTDA	0.908
WP_016361171.1	CGGGGGGGGGGS	1	SCGGGGGGGGGG	1
	PKPSTPSTPSNP	1	PKPSTPSTPSNP	1
	NQERALKGPGAF	1	EHGEEVLEVVRD	1
	EDKDGKTVDENG	0.999	NQERALKGPGAF	1
	QTLFTTTDKTDL	0.998	LGENAEASSKNV	1
	YAPNNYAETPGI	0.996	RRGSFDESEASW	1
	EHGEEVLEVVRD	0.989	TSTGDTVTGKIN	1
	TSTGDTVTGKIN	0.987	TKFSPTTTLGVA	0.999
	PTQEVYEAAMAK	0.986	DSEYKPYKGEGN	0.999
	KFSPTTTLGVAL	0.982	DESPNPPYTKEA	0.999
	TGNNSFSGGSKV	0.974	AGNKGDKNASVE	0.998
	WSAGNKGDKNAS	0.969	TTTDKTDLSQDP	0.998
	DLPYGFYTAGRL	0.957	TPGIDNKPNNKI	0.997
	LGENAEASSKNV	0.952	FDNDIYGDGGLE	0.997
	DESPNPPYTKEA	0.952	NTDKRDLSYEGR	0.995
	TNITRNPDGTIR	0.952	GANSVEDKDGKT	0.994
	KGTTAIIGGDYV	0.947	GAQEEGLRIRSG	0.994
	GIAIPKDTTEID	0.942	ENVEKVFEDLDQ	0.994
	SEYKPYKGEGNL	0.928	RITNITRNPDGT	0.993
			DLPYGFYTAGRL	0.992
			DKLSSAGFEARY	0.992
			HLTGNNSFSGGS	0.99
			GGDYVAYAGSNT	0.986
			NISYYGDTTIFK	0.985
			VGYDVDNFDLIN	0.977

			EDKKWADSVFSA	0.975
			RYAGESKSDMVG	0.971
			STMGFTSATEMM	0.97
			IPKDTTEIDGSG	0.967
			FPTVTNPLDSQK	0.96
			ASPITVGTSGTL	0.959
			GSSYAAPKVTRA	0.946
			DGYASADTRVTG	0.943
			DENGHPLVPRRS	0.913
			PYFDNKLEKGWI	0.904
WP_016361319.1	FFDFGGLAGGGK	1	EKEEGNFKEKYK	1
	NPAGTARFEKGK	0.992	GGLAGGGKLNYD	1
	AIKKQKTNEALT	0.974	GIYDKGSSLTGE	0.999
	SPDYLSNQAQTG	0.966	DSVSPFYNPAGT	0.995
	AQFKPLGIYDKG	0.943	GRAFSVNDKLSF	0.98
	VTLGAGVRYNFD	0.94	VNPSTEYKQAKA	0.973
	PSTEYKQAKAQQ	0.939	TYSPDYLSNQAQ	0.964
	FKEKYKVDENQK	0.928	TSTDDNGAYFFD	0.964
			GAGVRYNFDETL	0.964
			EHGRDYKDGWEV	0.949
			SFNDTEYALNSV	0.906
WP_017139894.1	AGAGNGTVGNNN	1	KGENNTTSGKSN	1
	AGAGNGTVGNNN	1	GGASQTTTGNNN	1
	AGAGNGTVGNNN	1	GAGNGTVGNNNK	1
	AGAGNGTVGNNN	1	GAGNGTVGNNNK	1
	AGAGNGTVGNNN	1	GAGNGTVGNNNK	1
	AGAGNGTVGNNN	1	GAGNGTVGNNNK	1
	AGAGNGTVGNNN	1	GAGNGTVGNNNK	1
	AGAGNGTVGNNN	1	GAGNGTVGNNNK	1
	AGAGNGTVGNNN	1	GAGNGTVGNNNK	1
	AGAGNGTVGNNN	1	GAGNGTVGNNNK	1
	AGAGNGTVGNNN	1	GAGNGTVGNNNK	1
	TTIGNNNTIGGS	1	GAGNGTVGNNNK	1
	AGAGNGTVGDNN	1	GAGNGTVGNNNK	1
	AGAGNGTVGDNN	1	HDNITSGDNSLA	1

AGAGNGTVGSNN	1	HDNITSGDNSLA	1
AGAGNGTVGSNN	1	AGAGNGTVGDNN	1
AGAGNGTVGSNN	1	AGAGNGTVGDNN	1
AGAGNGTVGSNN	1	TTIGNNNTIGGS	1
LKGENNTTSGKS	1	GANNGTVGNNNK	1
VAGAGNGTVGNK	1	GATNKTTGDKNN	1
VAGAGNGTVGNK	1	NKTTGNNNLILG	1
VAGAGNGTVGNK	1	TTTGNNNKIGGA	1
VVGGASQTTTGN	1	NGIGNKTTGIKN	1
GNNNKVIGNNNF	1	AGAGNGTVGNKN	1
TTTGNNNKIGGA	1	AGAGNGTVGNKN	1
GANNGTVGNNNK	0.999	AGAGNGTVGNKN	1
LNNTVKGNHNNV	0.999	AGAGNGTVGSNN	1
YGNDNTVNGNGS	0.999	AGAGNGTVGSNN	1
LSNKTTGNNNLI	0.998	AGAGNGTVGSNN	1
KTTGNGNNVLGT	0.998	AGAGNGTVGSNN	1
GNDNEVTGENSS	0.998	YGNDNTVNGNGS	1
GNDNEVTGENSS	0.998	GSFNKTTGNSNV	1
FNKTDGKNNYIV	0.997	TGASNTTKGDEN	0.999
GNNSGAFGDPNT	0.996	SGRSNETTGNNN	0.999
NKTTGDKNNITG	0.994	GANNGTVGDNNW	0.999
TGASNTTKGDEN	0.994	GANNGTVGDNNK	0.999
AGTDNGTVGDNN	0.994	AGTDNGTVGDNN	0.999
VAGAGNRTEGSK	0.993	EKIKNLEEKLEK	0.999
FGHDNITSGDNS	0.991	NKTTGNGNNVLG	0.999
FGHDNITSGDNS	0.991	NTNVGNSDITNG	0.998
KNKTTGDKNLVA	0.989	GDNNKVSGKDNL	0.998
AGAGNKTAGNSN	0.985	ENKELKSKVEVQ	0.998
AGAGNKTAGNSN	0.985	GVEYNETPQYVV	0.998
GANNGTVGDNNW	0.983	GNNSGAFGDPNT	0.998
AGANNGTVGSNN	0.982	DNEVTGENSSAF	0.998
AGANNGTVGSNN	0.982	DNEVTGENSSAF	0.998
AGANNGTVGSNN	0.982	ASNGTVGDNNKV	0.998
AGANNGTVGSNN	0.982	LNNTVKGNHNNV	0.998
AGANNGTVGDNN	0.98	GNGSVASTSNEV	0.997

SNETTGNNNIVS	0.979	VEGEINNVKDEV	0.997
AGASNGTVGDNN	0.976	NQNKVEGKESSA	0.995
RGNNNITDGRSN	0.975	AGAGNKTEGNKN	0.993
AGAGNRTEGYKN	0.969	KNKTTGDKNLVA	0.992
GLANKTTGDKNL	0.956	GANNGTVGSNNW	0.99
AGASNGTVGSNN	0.95	GANNGTVGSNNW	0.99
SNKTEGTSNEIL	0.949	GANNGTVGSNNW	0.99
ENKELKSKVEVQ	0.937	GANNGTVGSNNW	0.99
LGIGNTVKGSSS	0.928	AGDDNKVAGIGN	0.989
NGIGNKTTGIKN	0.921	AGDDNKVAGIGN	0.989
NTNVGNSDITNG	0.902	AGDDNKVAGIGN	0.989
		AGDDNKVAGIGN	0.989
		AGDDNKVAGIGN	0.989
		NGVANTTAGSRN	0.988
		GDGEYSAKSTDA	0.987
		AGAGNKTAGNSN	0.984
		AGAGNKTAGNSN	0.984
		IGDDNKVAGIGN	0.981
		AGAGNRTEGYKN	0.981
		LGRNNKVKGEDG	0.978
		GISNTLKGDSNK	0.977
		GRDNEAKGERNL	0.975
		GRDNEAKGERNL	0.975
		GKENETTGENSL	0.973
		SNKTEGTSNEIL	0.971
		ITFKSGVSNSVA	0.971
		QYDPKAPTQVMA	0.97
		IGNNNKVIGNNN	0.968
		GTFNKTEGNKNS	0.964
		GNNNITDGRSNE	0.958
		GASNGTVGSNNW	0.956
		VGEGNKTEGKYS	0.953
		VGEGNKTEGKYS	0.953
		LGRDNKVKGENS	0.952
		KGNIAKGNSSTA	0.949

			KGNIAKGNSSTA	0.949
			KENKTDGINSTA	0.946
			KENKTDGINSTA	0.946
			NGIGNKVAGAEN	0.943
			IADGRTNKITGD	0.938
			GLANKTTGDKNL	0.908
			VSGTENKVIGNK	0.9
WP_023038391.1	ARSVPGDPEGID	1	EARSVPGDPEGI	1
	NTPIGPLRFDFG	0.999	VVDVVEKENVVD	1
	GKFQVEGGHAGG	0.999	KFQVEGGHAGGY	1
	LTTKPGSVQNYN	0.999	NTLREDTDKSIV	1
	ESQKFWVGGGNS	0.994	QNYNNLREDRDK	0.998
	FGHNIGTTAGVG	0.988	WLWSVYPYISYD	0.998
	TLREDTDKSIVI	0.981	GGGNSLRGYDGG	0.998
	IVEGIVRKIEVK	0.973	VEGIVRKIEVKK	0.997
	RGYDGGFFKGSQ	0.957	FGHNIGTTAGVG	0.995
	LKDSNWRGKNQE	0.956	NRRTPNDDVLKT	0.99
	VYPYISYDTRNN	0.954	VLQPTSYDGGVR	0.989
	VGHKMDDDGMKF	0.943	VITGNHTIPTST	0.989
	NRRTPNDDVLKT	0.925	EVVNNQQVPASL	0.987
	EKSNKDYTGFAL	0.919	QGGVAYGSETGL	0.981
			RLNTPIGPLRFD	0.972
			NVKEYDATVDNL	0.969
			KTSYGDEDSELF	0.967
			TDMSTDENGTLH	0.957
			ADFNALKETGYF	0.949
			DSNWRGKNQEFG	0.902
WP_060675827.1	KRKNSTTTTIMV	0.996	KRKNSTTTTIMV	1
	TKEGRAQNRRVE	0.995	FDFDKSIVKAQY	1
	GIEAMGEEQPIA	0.909	YNFGLSRRRAEA	0.995
	DKSIVKAQYFDL	0.909	DILNSEAPKEMT	0.964
			MGEEQPIATNAT	0.951
			LTTTQMRENSIR	0.913
WP_060676683.1	KEVMPAPTPAPE	1	KEVMPAPTPAPE	1

	GGEGRDYAPYDG	0.999	WTGHSNGGEGRD	1
	GTKGDASPDEYR	0.999	ETENKTKSEDTD	1
	SKGNKKGQYYGA	0.979	APYDGYDAGTNN	1
	DREVAPAWRPNG	0.967	RGTKGDASPDEY	0.999
	ISRLEYKQKDGS	0.965	VVEYVEKPVIVY	0.999
	EDTDRDWARSRN	0.962	PAWRPNGSVDVQ	0.998
	FEGGYDPYTWHQ	0.948	LSFDFEGGYDPY	0.986
	VYNGYDRYQGSF	0.941	TTANVNFTKNQT	0.97
	KNQTLDVRVRNY	0.937	AKHAEASVLFDF	0.961
	KFGLRPGYAHNW	0.914	SRLEYKQKDGSE	0.953
	YQHTPLYKNNAV	0.903	YQHTPLYKNNAV	0.947
			LNFESEYTLPLG	0.946
			NVYNGYDRYQGS	0.926

Table 4. B-cell epitopes prediction

B-cell epitope screening for the 16 outer membrane and extracellular proteins are predicted by BCPREDS. BCPred represents the result of fix length prediction method and FBCPred is for the flexible length prediction, only non-overlapping epitopes were reported. Protein WP_008701556.1 and WP_008701556.1 contained much more epitope.

Table 5 T-cell Epitopes

Accession	HLA - A2 epitope	P value	HLA – A3 epitope	P value	HLA-DRB1*0402	P value
WP_016361171.1	MMIALASILFV	0.00308	SVFSAGLQY	0.000196	GGYFVQFNI	0.00161
	SNKDSEVWFSV	0.00345	KVKNYGTVK	0.0012	DRYYYWDGN	0.00161
	VLMGATVQSM	0.00441	VLSNGYVTK	0.00185	YKLNRRFFN	0.0031
	NLPINPGTPSV	0.00531	KVFEDLDQK	0.00293	TKSNVSFGI	0.00424
	AILDANFVNAV	0.00531	GAFMNISYY	0.00352	TNANAKNNI	0.00575
	TLTDVYTDV	0.00561	ILGSVGWDY	0.00421	KDYNKVESC	0.00619
	KFSPTTTLGV	0.00716	EVYEAAMAK	0.00472	NWGNYSKNK	0.00718
	KMMMIALASI	0.00898	QVGFKNSVK	0.00499	VVGFEMNSS	0.00718
	LLTSTGDTV	0.00901	GLRIRSGQY	0.00557	QHGRAAFVN	0.00829
	KLQAYVTQA	0.00954	TTLGVALNY	0.00724	KDYFVITPN	0.00956
	WMTADQIRQTL	0.0106	GWFTPFIGY	0.00724	GVYLAGANW	0.00956
	MMMIALASI	0.0107	KFYGVKQAK	0.00762	DDDNARLRW	0.00956
	KLMIPYFRDA	0.0124	IIGGDYVAY	0.00926	NYGVRLNYK	0.0103
	SMSTMGFTSAT	0.0133	GTIRSISSK	0.0102	TPSDNKNSI	0.011
	FLVGARAEYV	0.0138	MVGVSLYGK	0.0102	IPYNSLISG	0.011
	GLEHGEEVLEV	0.0145	SAYVELGKK	0.0139	TRNVDKFSV	0.0118
	KMLSAYVEL	0.0147	EITPVFGVY	0.0145	YDYDINTNW	0.0118
	KIKIQNGTVSV	0.0151	KSYDDSEYK	0.0171	VYGRQGILN	0.0126
	ALKGPGAFMNI	0.0165	GIDNKPNNK	0.0185	IGIYAMNNN	0.0126
	SNTQVGFKNSV	0.0187	LINNVDAQK	0.0185	KDSVVDFNI	0.0135
	MMSGEIYASA	0.0204	ASVEAGLPY	0.0201	GIYYEGTGN	0.0135
	KLGLGTLHLT	0.0204	LFIWSAGNK	0.0201	QVNVTSGTI	0.0145
	AILPTQEVYEA	0.0212	TWIGFGAFR	0.0217	NMGYGQFSN	0.0154
	KTVDENGHPLV	0.0221	VTRAAALVY	0.0233	DKGVVTINN	0.0165
	KLMIPYFRDAV	0.0239		0.0242	GNINVGNNS	0.0176
	VLNPKAIVGY	0.0245		0.0251	NPNNTTMAA	0.0176
	FGAFREITPV	0.0245		0.0261	GKYYGEYGI	0.0176
	RLGLSHISSKV	0.0269		0.028	NVNFVKNAV	0.02
	ILDANFVNAV	0.0281		0.028		0.0228
	GLSHISSKV	0.0291		0.0357		0.0242
	GSVGWDYTYV	0.0294		0.0369		0.0258
	GMLNQERAL	0.0304		0.0369		0.0258
	VMADGTEGV	0.0319		0.0407		0.0258
	GVYGNVDFRV	0.0321		0.0421		0.0258
		0.0326		0.0449		0.0258
		0.0333		0.0463		0.0274
		0.0335		0.0494		0.0274
		0.0349				0.0274
		0.0349				0.0291

	TKFSPTTTLGV	0.0349	QTLFTTTDK		GILNVKGAL	0.0291
	AYAGSNTQV	0.0351	MLNQERALK		VRGKRAVGI	0.0291
	AMAKFGNQSV	0.0351	AEYVGNKYK		NIKNTGTLN	0.0291
	GAFREITPV	0.0365	TYVENGQNK		GDSKGIFSM	0.0309
	ALVYDKFSWM	0.0378	KVHSGDSEK		NNHRTKRYS	0.0309
	KIQNGTVSV	0.038	GRFTGSSVK		VYSNIQERM	0.0309
	TLSRQNPSEYL	0.0381	TSYFDNDIY		GKDFTITSS	0.0328
	KWADSVFSAGL	0.0392	LSRQNPSEY		GKNNVAAYI	0.0328
	PINPGTPSV	0.0397	RLSGLDNFK		KVSVTKKAI	0.0328
	SFDESEASWGI	0.0413	DLEYAPNNY		TEGNVAGNL	0.0348
	VTGTERPTGL	0.0416	SWGIKADKK		GVNQAKMRN	0.0348
	KLEKGWISV	0.0416	ALKASLYTA		VDGNDLTST	0.0348
	EMMSGEIYASA	0.0416	KLRRDGYAS		KDGYQYFNR	0.0348
	GNLPLPFAKV	0.0416			QKHYIKVEN	0.0348
	LLTSTGDTVT	0.0421			TINMNRTTL	0.0348
	LMIPYFRDA	0.0434			GRGWIYSNS	0.0369
	SVGWDYTYV	0.0434			GKYGWSAGI	0.0369
	TVDENGHPL	0.0449			NGANAKFEN	0.0369
	GMRTANVEV	0.0454			TDGNLKLSG	0.0391
	NLPLPFAKVHS	0.0467			PKAVTEKTI	0.0391
	DLSNRLSGL	0.0473			VGIYAKNTT	0.0391
	TKFSPTTTL	0.0473			GVYAVNGSK	0.0414
	SLYGKQDLPY	0.0473			GNKITLKNI	0.0414
	KLSSAGFEA	0.0494			DKSGTKDDT	0.0438
	KKFGWFTPFI				QVGVAVKGT	0.0438
	NLPLPFAKV				LKGNVMAGT	0.0438
	YAAPKVTRA				TVKNGATGI	0.0438
	LVYDKFSWM				VPLRTEIKT	0.0438
	KFGWFTPFI				NYNYKIGGS	0.0463
					NSNMNMTLG	0.0463
					DNANIAQSG	0.0463
						0.0489
						0.0489
						0.0489
						0.0489

					GMAYREDAS GKNTAATLN TNTYTMKLG GIYLAGTAS ISYGHIFNI DFARAVANI GIYTTNDNS SKNNIINTL GALNVADST YVDNTATSK QDGGTITTG QSNVTLNEN TPPNEPNPS GVYVVDGEA GIYGANGGS GYNKTVNTD NMDLTLNNF SDDVIAFNL NIYLFAQGG LPSFSRGNS	
A0A140PUD9	YLFSNNFFKV KNWRWQPTAWA SLLVVGSVASA ELYMLPTFQV LNFESEYTLPL YLFSNNFFK ALVLGSLLVV YLYQHTPLYK FLNFESEYTL LVLGSLLVV NVSRLQTTANV YTLPLGFSA	0.000191 0.00454 0.0056 0.006 0.00686 0.00804 0.00898 0.0124 0.0131 0.0163 0.0239 0.0253 0.0257	YLFSNNFFK YLYQHTPLY KVIANGSGY TTANVNFTK NVYNGYDRY RWYGETENK TLDVRVRNY AVTNNSEAK	0.000317 0.000955 0.00161 0.00293 0.00312 0.00588 0.00801 0.0133 0.0185 0.0225 0.0251 0.0333 0.0369	APAWRPNGS YRNFAVTNN NWRWQPTAW DVQYRWYGE DWANSKNNV WHQYKVIAN DGYNGGTNN ADYLFSNNF	0.00458 0.00619 0.00718 0.00772 0.00829 0.0118 0.02 0.0228 0.0258 0.0328 0.0348

	RWQPTAWAGM KNWRWQPTA YLYQHTPLY LYMLPTFQV VLFDFADYLFS	0.0278 0.0304 0.0319 0.0326	SYKPTDFVK YVEKPVIVY TSRLEYKQK DFEGGYDPY GLRPGYAHN REYVTFDGY GTNNKYFLN VVGSVASAK	0.0407 0.0449 0.0478	GYDRYQGSF NWSGHSNGG IRHFYNFGN	
A0A140PSK0	SILMLFAFV ALENNFWKVSV MLFAFVACSST ALENNFWKV FVACSSTQFV KTMDSIPGASI IFLSILMLFA VLIQYFPTEF YLGTYQGII QYFPTEFEIAL VLIQYFPTE	0.000765 0.00345 0.00478 0.00674 0.00803 0.0212 0.0234 0.0365 0.0398 0.0467 0.0494	TCEGIGADY KALENNFWK KRIAVYKFK SATVTQAVK	0.00926 0.0102 0.0164 0.0494 0.000765	NIIKTMDSI LKSNFAITC DYLGTYQGI	0.00891 0.0214 0.0438
A0A140PPP98	LLLALLVTA KLYAVLLLAL ILFDFDKYIV LLLALLVTACS NLILSMPELIL SLSTLAKAL NLILSMPEL ILFDFDKYI FIGTEAYNLEL	0.00252 0.00469 0.00637 0.00652 0.0158 0.0253 0.0253 0.0304 0.0378	FDFDKYIVK SVKRANAIK LVTACSSTK AVEDTEANK KLYAVLLLA	0.00801 0.00971 0.0133 0.0164 0.0185	DKYIVKDKI GEQNPADTN SRNRRVEFI TKTNRKTNV MPELILFDF VTIRREGNN	0.00829 0.0135 0.0135 0.02 0.02 0.0258
WP_008701980.1	TTLPSEPTIHI LLGDRSMGWV FLSGSKVTV MLVDKPLSV VIMGWAYKHL	0.000628 0.00124 0.00149 0.00174 0.00259	FLINGGFSY ALYASGTTK TLNNVIAYY	3.37E-05 0.000263 0.000816 0.00129 0.00226	GGYFVQFNI DRYYYWDGN YKLNRRFFN TKSNVSFGI	0.00161 0.00161 0.0031 0.00424 0.00575

	KLGLHYQAPL	0.00277	GVYAVNGSK	0.00242	TNANAKNNI	0.00619
	TLTAPEEVAIV	0.00308	GIYDKSGTY	0.00258	KDYNKVESC	0.00718
	GLIDNQGTINV	0.00326	ASYGKNMGY	0.00331	NWGNYSKNK	0.00718
	GLYGNKGTV	0.00357	ALNGGGTLK	0.00331	VVGFEMNSS	0.00829
	NLLGNKVTL	0.00382	SLISGERYK	0.00397	QHGRAAFVN	0.00956
	YLAGANWER	0.00382	GLLVKGTGK	0.00421	KDYFVITPN	0.00956
	GIYAKNTTV	0.00435	SVIMGWAYK	0.00472	GVYLAGANW	0.0103
	YLSGNSTAA	0.00435	GVINAQKHY	0.00472	DDDNARLRW	0.011
	GLFGTKKIGL	0.00469	GAIGAFINY	0.00557	NYGVRLNYK	0.011
	TMVRIPYNSLI	0.00478	IVEKTVIKK	0.00588	TPSDNKNSI	0.0118
	VMAGTNVNFV	0.005	TLGVSGYEY	0.00653	IPYNSLISG	0.0118
	KTLTSTGNITV	0.00504	AVYSDGTGK	0.00653	TRNVDKFSV	0.0126
	LINNMNVTV	0.00527	NILKGYQSY	0.00688	YDYDINTNW	0.0126
	SIINSNMNMTL	0.0056	GVFTDGNLK	0.00762	VYGRQGILN	0.0135
	QLTDTNATGQV	0.0056	TAEGRGWIY	0.00762	IGIYAMNNN	0.0135
	YLTKPIDLL	0.00561	SISNNFDIK	0.00801	KDSVVDFNI	0.0145
	YLNDREAFT	0.00561	GLKVDIINK	0.00801	GIYYEGTGN	0.0154
	ALGSIGKTLTV	0.00589	GIETSIKSK	0.00801	QVNVTSGTI	0.0165
	FLTLKNANSSI	0.0062	GVTINYNYK	0.00801	NMGYGQFSN	0.0176
	KLTLESPNGHI	0.00652	FVKNAVPLY	0.00841	DKGVVTINN	0.0176
	TMAAGSSVYLT	0.00652	KTAVGIYAK	0.00841	GNINVGNNS	0.0176
	FLAQRLKATA	0.00716	KLNSGTKIK	0.00926	NPNNTTMAA	0.02
	GLVNIDATSKI	0.00721	ASRSSILAY	0.00926	GKYYGEYGI	0.0228
	ITVQSGSTITV	0.00757	ILDSNNQEK	0.00971	NVNFVKNAV	0.0242
	TINMNRTTL	0.00759	GMYLDTAVK	0.00971	GILNVKGAL	0.0258
	FMADDTNPNNT	0.00795	YLAGTASLK	0.0107	VRGKRAVGI	0.0258
	GLGSGSLNNL	0.00849		0.0107	NIKNTGTLN	0.0258
	ITLESGSTLNV	0.00876		0.0112	GDSKGIFSM	0.0258
	TLLGGGANPQA	0.00876		0.0112	NNHRTKRYS	0.0258
	KLTQRGEFWV	0.00898		0.0112	VYSNIQERM	0.0274
	YDINTNWTV	0.00954		0.0112		0.0274
	WLTRGEITV	0.0101		0.0117		0.0274
	FLINGGFSYA	0.0106		0.0117		0.0291
	MLVDKPLSVGI	0.0111		0.0117		0.0291
	TVADKGVGLLV	0.0121		0.0122		0.0291
	RLSYSMALL	0.0125		0.0128		0.0291
	LQHGRAAFV	0.0125		0.0128		0.0309
	NLHGSSTAFDV	0.0127		0.0133		0.0309
		0.0127		0.0133		0.0309
		0.0131		0.0139		0.0328
		0.0133		0.0145		0.0328
		0.0139		0.0151		0.0328
		0.0145				

	TTLNNAGEINL	0.0145	KMRNASTGY	0.0158	GKDFTITSS	0.0348		
	LLGGGANPQA	0.0145	FIFANWGNY	0.0158	GKNNVAAYI	0.0348		
	GILADRGTSSI	0.0145	KTVENSFDK	0.0164	KVSVTKKAI	0.0348		
	TLNNFHVSA	0.0145	SVGILGMAY	0.0171	TEGNVAGNL	0.0348		
	KLTFDLATGKL	0.0145	TQRGEFWVK	0.0178	GVNQAKMRN	0.0348		
	NNTDTTVTSHL	0.0147	MLANAASGK	0.0178	VDGNDLTST	0.0369		
	IALNGGGTLKV	0.016	SITGNVTNY	0.0193	KDGYQYFNR	0.0369		
	DLLGDRSMGWV	0.016	KLEYTGNGY	0.0193	QKHYIKVEN	0.0369		
	KLNRRFFNTL	0.016	IAEKGYGLY	0.0201	TINMNRTTL	0.0369		
	ALQGITVAAV	0.0165	GQFSNGKYY	0.0201	GRGWIYSNS	0.0391		
	GLIDNQGTI	0.0171	IVGSNFEFK	0.0201	GKYGWSAGI	0.0391		
	NLLALYANGL	0.0172	QAKKDGYQY	0.0208	NGANAKFEN	0.0414		
	TLPSEPTIHI	0.018	IVASKDGTK	0.0217	TDGNLKLSG	0.0414		
	FIAQNTTFNV	0.0189	SVAKVSELK	0.0225	PKAVTEKTI	0.0438		
	TLLNVPYSEFS	0.0195	ITVNNHRTK	0.0225	VGIYAKNTT	0.0438		
	NIQERMKTV	0.0199	TFRAEYQGY	0.0233	GVYAVNGSK	0.0438		
	FTITSSSTNKV	0.0209	KISGGTGQK	0.0233	GNKITLKNI	0.0438		
	HLVTNKGTV	0.0212	VIYTGGEHK	0.0251	DKSGTKDDT	0.0438		
	GLTVAGVSV	0.0212	GLVNITNEK	0.0251	QVGVAVKGT	0.0463		
	TLKNIFNVSPS	0.0214	GTSAAGIYY	0.0261	LKGNVMAGT	0.0463		
	ITLKNIFNV	0.0214	ATELLFINK	0.0261	TVKNGATGI	0.0463		
	TTTSDITMV	0.0214	GAVHQAFAK	0.0261	VPLRTEIKT	0.0489		
	KLTQRGEFWVK	0.0221	YVDNTATSK	0.027	NYNYKIGGS	0.0489		
	SFANDNFNGQV	0.0224	GANAKFENK	0.027	NSNMNMTLG	0.0489		
	LLVKGTGKTL	0.023	GSNNSVGLY	0.028	DNANIAQSG	0.0489		
	SLGTNAKGLV	0.023	SSTGLFGTK	0.028	GMAYREDAS	0.0489		
	IMADRTDAGA	0.0234		0.029	GKNTAATLN	0.0489		
	KFIAQNTTFNV	0.0239		0.029	TNTYTMKLG	0.0489		
	VLPLGAVHQA	0.0239		0.029	GIYLAGTAS	0.0489		
	LLFINKGNINL	0.0239		0.029	ISYGHIFNI	0.0489		
	MTPPNEPNPSV	0.0241		0.03	DFARAVANI	0.0489		
	KLINNMNVTV	0.0241		0.03	GIYTTNDNS	0.0489		
	GTMNINASSPI	0.0245		0.03				
	NVMAGTNVNFV	0.0245		0.0311				
	TIGSSGFTTPS	0.0253		0.0311				
	FLKENSIGL	0.0253		0.0311				
	GIYSQGGTV	0.0253		0.0322				
	ILNITAPTA	0.0253		0.0322				
	KLNSGTKIKV	0.0253		0.0322				
		0.0265		0.0333				
				0.0333				

ALNGGGTLKV	0.0265	TEFNYDVTK	0.0333	SKNNIINTL
YITGNSTFR	0.0269	PLEVKGKDY	0.0333	GALNVADST
KLGKSSTGL	0.0269	GVSAGNGGY	0.0345	YVDNTATSK
TLNENSIGI	0.0269	AFVNAGEIK	0.0357	QDGGTITTG
ALYANGLKL	0.0269	NVSFGIAAY	0.0369	QSNVTLNEN
ALNMEGLGL	0.0269	IVGKITGTK	0.0369	TPPNEPNPS
LLYGDGTNT	0.0269	DVYSNIQER	0.0369	GVYVVDGEA
ALQGITVAA	0.0269	TMAAGSSVY	0.0394	GIYGANGGS
SVTPPTVEV	0.0278	YLAGANWER	0.0394	GYNKTVNTD
ITANATTTV	0.0278	GILQDHSSK	0.0394	NMDLTLNNF
EFLSGSKVTV	0.0278	GTAPTSGTY	0.0394	SDDVIAFNL
KRLSYSMALLI	0.028	DVIAFNLKR	0.0407	NIYLFAQGG
KISAENIALV	0.0281	ALYANGLKL	0.0407	LPSFSRGNS
NYLDNEGTISA	0.0281	STHEKGNIY	0.0407	
TMTIAEKGYGL	0.0281	SLDNISSHR	0.0407	
DMSDGNAKVSV	0.0281	FSRGNSEQK	0.0421	
SVVDFNIAPTV	0.0294	NIENATGAK	0.0421	
LTQRGEFWV	0.0294	EAFTYGGKY	0.0421	
YATGEHSTI	0.0302	SLGLYAKDK	0.0434	
GLFAKDGGV	0.0302	ATANGSTIK	0.0449	
KMYADGEGSVI	0.0304	KSAEGIGMY	0.0449	
YFVITPNVGV	0.0307	SGIGLYGNK	0.0449	
RLRWLTRGEI	0.0307	IIETTSGEK	0.0449	
RLITPPEVPA	0.0307	TVTVEGTHK	0.0449	
GLSVGMLANA	0.0314	KVESCLKSF	0.0463	
GIYNETTGAV	0.0314	KGKNNVAAY	0.0463	
ITGSGSFLTL	0.0319		0.0463	
NMDLTLNNFHV	0.0326		0.0478	
TVVDAKSISGV	0.0326		0.0478	
FVITPNVGV	0.0333		0.0478	
TMAAGSSVYL	0.0335		0.0494	
LNMEGLGLDV	0.0335		0.0494	
AMNNNTDTTV	0.0349			
KITANATTTV	0.0349			
GMLANAASGKT	0.0349			
QITSATGKAHI	0.0349			
GMYAKGGAL	0.0349			
FLINGGFSYAD	0.0349			
NLKRATGLTTV	0.0351			
	0.0351			

KLNRRFFNT	0.0351	KLNRRFFNT
YLSGNSTAAY	0.0351	KLVGAYFKD
YLKGSGSNLL	0.0364	QNSNKFEFY
AIYGKNTAA	0.0364	GHIFNIKEK
DLTSTGTVTV	0.0364	KKAIGIYAK
YVFENITNA	0.0365	KVINGGILN
GLYVDNTAT	0.0365	GNNSVGIYK
FLAQRLKAT	0.0365	SGLEDIRSK
YLKGGEVDL	0.0378	STGNATIGK
FIVGKITGT	0.0378	KEAVGIYAK
YIFTTTHTG	0.0378	NSIGLYGAK
GTLNLTKTGAV	0.038	LHYQAPLNK
AVINLETEGNV	0.038	ANNNGIGVY
TNVDETAINLV	0.038	TTVEVDGVK
SLNNLFSGSTA	0.038	STAAYAKDK
TIKDVFRTGVV	0.0381	DFKSGIGVY
KLYHGVNQAKM	0.0381	AGTGMYLTK
INANGEVIGMV	0.0381	GNLSTAVTK
LTSTGTVTV	0.0392	DASHNPIVK
YIPEHSKFI	0.0392	HSKGSIGVY
KLSGNGGLI	0.0392	AFSGDVKSK
GILNVKGALNM	0.0397	ASAELGGIK
ALASVAKVSEL	0.0398	GKDSIGVYK
ITAPSVTPPTV	0.0398	SYNELLSSY
YLDNEGTISA	0.0413	TGNSTFRNK
GLYGAKGATI	0.0413	NVGSGIGAY
LTFDLATGKL	0.0416	
IINSNMNMTL	0.0416	
KLYHGVNQA	0.0421	
ALGSIGKTL	0.0431	
NVTPSTPEV	0.0434	
VVTAGENLLAL	0.0434	
LITPPEVPADI	0.0449	
FVDKIDFGANI	0.0449	
GMYLTKGSHL	0.0449	
VINGGILNV	0.0449	
GIYAKGVNV	0.0451	
GIFSMGAGNV	0.0451	
ATGEHSTITV	0.0454	
FINKGNINL	0.0454	
	0.0467	

	SIIKTEVAV	0.0467	QSGSTITVK
	IAQSGNITV	0.0467	TLTVADSGY
	NVTPSTPEVAA	0.0467	ITAPTAKDK
	TAPSVTPPTV	0.0473	SIINSNMNM
	ITLTAPEEV	0.0473	LSWKNIISY
	ITLESGSTL	0.0484	GNLDKSGTK
	TLESGSTLNV	0.0484	EKENNVVNK
	TLNLTKTGAV	0.0486	KTALGSIGK
	ALYASGTTKI	0.0486	TITSSSTNK
	KLGGDKGTGI	0.0486	GKGIGLFAK
	TLKVSSAGQGT	0.0494	ESNFNGPEY
	AMTVGKDSIGV		VYTKSGYNK
	SLGTNAKGL		KGKESVIIK
	TINDKTASM		NLISEKGIY
	LTSTGNITV		KSTAKPDTK
	ELNGENSVGA		TLTQSKDSK
	ALGSIGKTLT		
	LLYGDGTNTY		
	YIFTTTHTGG		
	HLSGGKLEL		
	GIYAEDSQV		
	AVTNNGALNVA		
	TINAPTVTPPA		
	YYWDGNDGAI		
	ILADRGTSSI		
	LTAPEEVAIV		
	DMSDGNAKV		

Table 5. The T-cell epitopes prediction of the top five abundant proteins

T-cell epitopes are predicted against HLA-A2, HLA-A3 and HLA-DRB1*0401 alleles individually. The p-value cut off was assigned as 0.05 for a balanced binding probability and specificity. The length of peptide against HLA-A2 ranged from 9 to 11, while for HLA-A3 the length was consistent with 9 residues. Among the five selected proteins autotransporter proteins (WP_008701980.1) had most predicted T-cell epitopes follow by the most abundant autotransporter protein (WP_016361171.1) with 188 and 63 HLA-A2 peptides respectively.

Metabolic pathways and putative function

The metabolic pathways of proteins from *F. nucleatum* OMVs were classified by KEGG based on previously identified protein functions, and 45 of 98 proteins were annotated. Thirty four percent were classified as proteins that participated in genetic information processing, while 18% were involved in environmental information processing, and 9% in cellular processes (Figure 14). Eighty five of the 98 proteins were functionally annotated by EggNOG 4.5.1 database, and 11 groups were assigned (Figure 15. In spite of their unknown functions, 20 proteins were mapped to translation, ribosomal structure and biogenesis, and 12 were placed in cell membrane biogenesis (Figure 15). Most of the autotransporter proteins were classified as unknown, though the most abundant one, WP_016361171.1, was identified as being involved in post-translational modification, protein turnover, and chaperones.

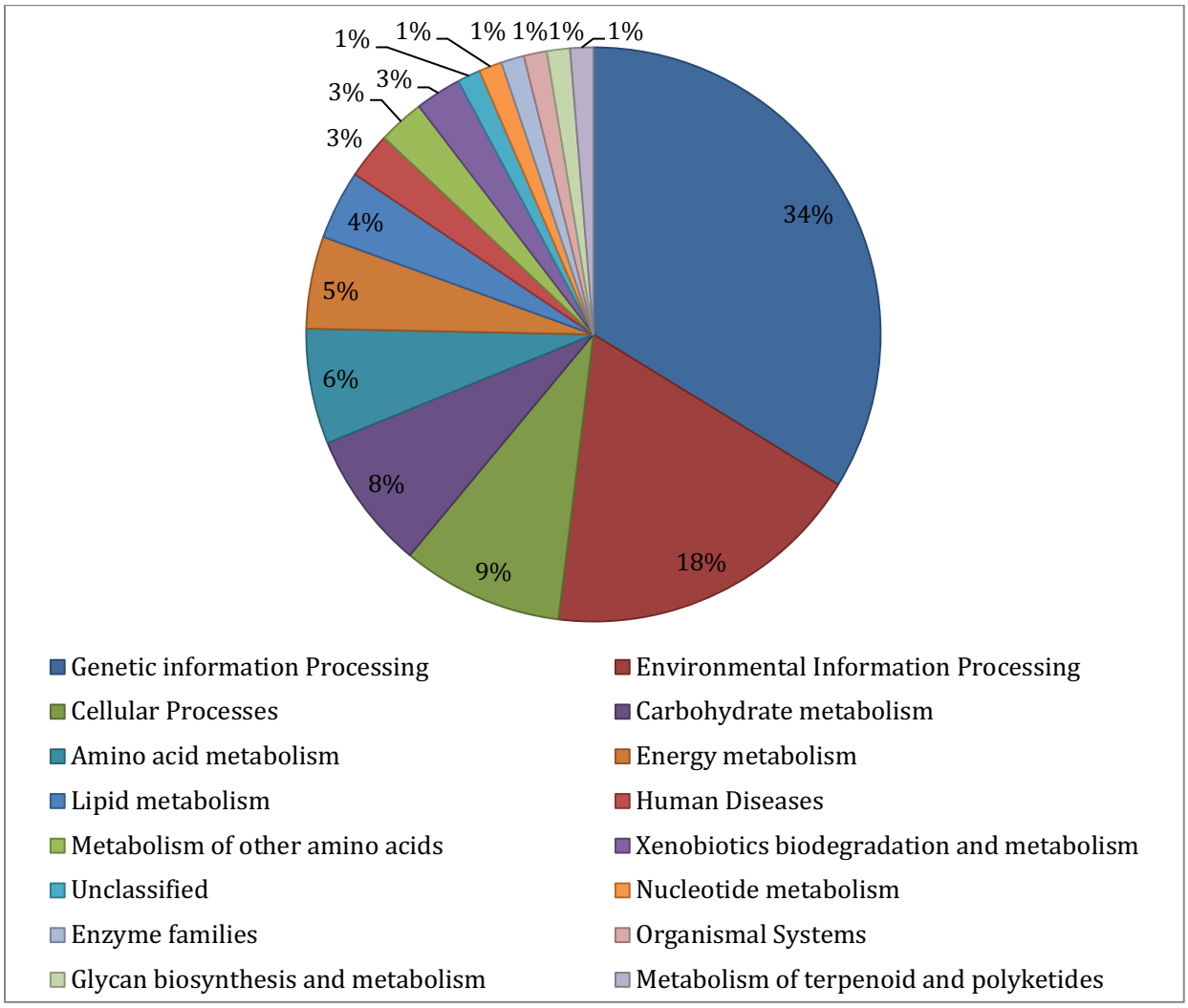


Figure 14 Metabolic pathway

45 proteins are annotated in KEGG database. 34% of them are found evolving in genetic information process. 18% participate environmental information process, 9% contribute to cellular process, and 8% play role in carbohydrate metabolism. Subsequently, small portion mediate the amino acid metabolism (6%), energy metabolism (5%), lipid metabolism (4%), human disease (3%), metabolism of other amino acid (3%), and xenobiotic biodegradation (3%).

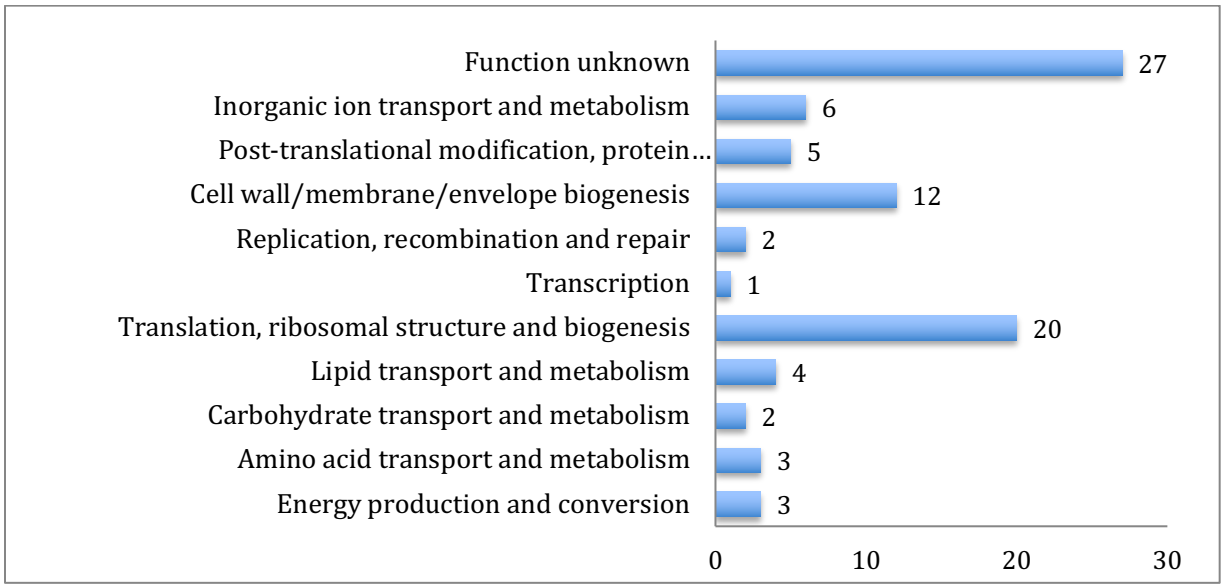


Figure 15 Function prediction of 85 protein annotated by EggNOG 4.5.1 database

The protein function of *F. ncleatum* is not well characterized. The function for 27 out of the 85 annotated proteins is unknown. In spite of unknown function, 20 proteins are packaged to translation, ribosomal structure and biogenesis, and 12 are mapped to cell wall/membrane/envelope biogenesis. Multiple proteins also mediate inorganic ion transport and metabolism (6), post-translational modification, protein turn over (5), lipid transport and metabolism (4), amino acid transport and metabolism (3), energy production and conversion (3).

Discussion

Protein components of *F. nucleatum* 7_1 OMVs

The OMV proteome of *F. nucleatum* has identified 157 and 190 proteins from the duplication experiment mass spectrometry analysis. Ninety-eight proteins were commonly identified in both OMV batches; 60 were predicted as secreted proteins with signal peptides. Among these secreted proteins, 16 were anticipated to be located in the outer membrane fractions, including extracellular proteins, which were the class of proteins most commonly seen in the OMVs. Outer membrane proteins consist of over half of the OMV protein residue totals, with others from cytosol, inner membrane, and periplasm fractions (Figure 12). By localization prediction of the 98 OMV proteins, cytosol proteins consist of half of the total, followed by the outer membrane, which is consistent with the previous findings [38, 63]. However, other studies with different Gram-negative bacteria indicated outer membrane (and extracellular) proteins were the major component of OMVs [64-65]. This discrepancy was not likely to be caused by contamination as the applied purification methods were virtually identical. Rather, varying results may be due to differences in biochemical profiles across bacterial species and culture conditions [65-66]. Some of the proteins found in OMVs may be secreted through non-classical pathways that are signal peptide independent. In this study, five cytosol proteins were predicted as being processed via non-classical secretion pathway; these had no predicted signal peptides (Table 2) [53]. The overall variety of OMV proteins notwithstanding, it appears that outer membrane proteins are the predominant class carried by the OMVs.

Type V secretion system

Type V secretion is also called autotransporter secretion, as there is no net energy consumption in this Sec-dependent secretion system [67]. Of all the 98 proteins reproducibly identified within *F.*

nucleatum OMVs, six autotransporter proteins were recognized and affirmed as the major component due to their copiousness, making up 31%~51% of the total protein load. Autotransporter proteins consist of three major domains: (1) the signal sequence at N-terminus that exports the protein through inner membrane; (2) the surface presenting passenger domain; and (3) the C-terminus β-barrel domain containing an autotransporter that modulates passenger domain exposure to surface [68-69]. Most recognized autotransporter proteins are known or predicted virulence factors; the diversity in passenger domains of autotransporter proteins shows multiple virulence-associated functions, including adhesion and cytotoxin [69]. For example, Fap2, an autotransporter protein of *F. nucleatum*, participates in multiple pathogenesis pathways, including binding to colorectal cancer cells, inducing lymphocyte apoptosis, and interacting with T cell immunoreceptors with Ig and ITIM domains (TIGIT) to facilitate immune evasion [21, 70-72]. Functional sequence of the Fap2 is located on the passenger domain. Interestingly, the passenger domains of the six enriched autotransporter proteins surveyed in this study are heterogeneous. This suggests versatile pathogenic functions. In contrast, the C-terminal translocator domains are highly conserved (Figure 16) [68, 72]. Sequence analysis demonstrated that four out of the six proteins (WP_008701980, WP_016361366.1, WP_016361171 and WP_020788899) had a high similarity and contained conserved domains of AidA, an adhesion protein involved in diffuse adhesion at the passenger portion [73]. AidA was identified as a versatile virulence factor in *Escherichia coli*, mediating adhesion, autoaggregation, biofilm formation, and invasion [74]. After travelling across the outer membrane, AidA domain stays tethered to the surface and self-cleaves via autocatalytic action [75-76]. The enrichment of autotransporter protein suggests that *F. nucleatum* OMVs may play a vital role in tissue colonization and invasion. Genetically modified autotransporters have been developed as platforms to expose multiple polypeptides on the OMV surface as antigens and to

explore their potential utility as vaccines [77]. OMVs loaded with known antigenic autotransporter antigenic domains and epitopes may prove to be powerful tools for inducing effective immune responses in the host.

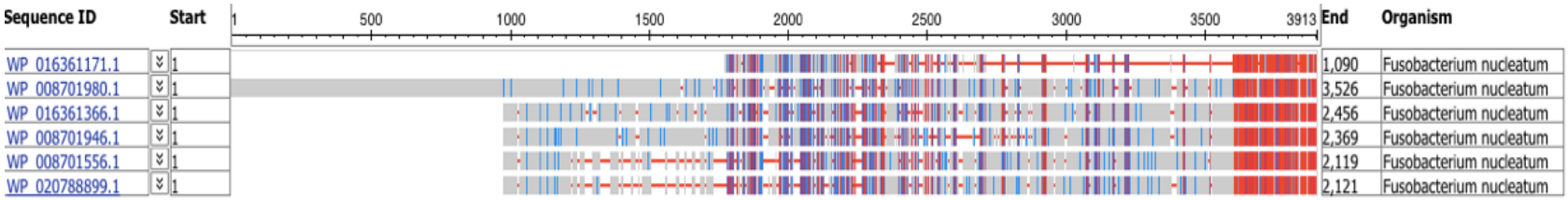

Sequence ID
Start
1
500
1000
1500
2000
2500
3000
3500
3913
End
Organism
WP_016361171.1 1 1,090 Fusobacterium nucleatum
WP_008701980.1 1 3,526 Fusobacterium nucleatum
WP_016361366.1 1 2,456 Fusobacterium nucleatum
WP_008701946.1 1 2,369 Fusobacterium nucleatum
WP_008701556.1 1 2,119 Fusobacterium nucleatum
WP_020788899.1 1 2,121 Fusobacterium nucleatum

Figure 16 Sequence similarity of six autotransporter proteins

The sequences of autotransporter proteins function domain are compared by BLAST. It reveals the C-terminal translocator domains are highly conserved and the passenger domain of the six enriched autotransporter protein are heterogeneous

Putative virulence factor carried by OMVs and potential application

Adhesion capability is crucial for bacterial pathogenesis, in particular for colonization and dissemination. Adhesins mediate this activity, and thus are major virulence factors. However, domains of adhesins that are exposed for immune recognition can be exploited if used as potential antigenic components of vaccines, particularly those known to be bound to or contained within OMVs. For vaccine delivery, antigen adherence and uptake are the first steps in eliciting immune response [34]. *F. nucleatum* FadA is one highly antigenic virulence factor (among others), and has garnered attention as an effective immunogen, which can be incorporated into OMVs as a vaccine against *F. nucleatum*. Previous studies have shown that FadA plays crucial role in *F. nucleatum* adhesion and invasion via binding to cadherin domains exposed on the host cell surface [26, 78]. Two FadA proteins (WP_005910368.1 and A0A140PS00) were found to be incorporated into OMVs isolated from *F. nucleatum*. The sequence of A0A140PS00 is highly conserved with a previous characterized FadA gene (FN0264), suggesting that OMVs may have similar immunogenic characteristics [29]. Another protein, one containing a MORN2 domain was found on OMVs. Previous studies analyzed by comparing whole-genome sequences suggested that proteins containing MORN2 may be involved in adhesion process, although the exact function has not been elucidated [79]. Two more cell surface proteins identified were predicted to have key bacterial virulence factors and contain a YadA (Yersinia adhesin)-like domain with possible immunogenic properties. This domain was characterized on *Yersinia* as a crucial virulence factor mainly involved in adhesion and combating host defense by resisting serum killing activity and phagocytosis [80]. All these putative virulence factors carried by *F. nucleatum* and other bacterial OMVs strongly suggest that these contain highly immunogenic domains and could prove useful if

incorporated into an OMV vaccine construct, as whole proteins or as recombinant fragments thereof.

Furthermore, B-cell and T-cell epitope predictions give researchers an opportunity to tailor OMV preparations that can incorporate antigenic domains from known virulence factors that elicit B and T cell responses. Exposed B-cell epitopes identified on OMV cargo proteins enhance the ability to induce effective antibodies against *F. nucleatum*. In accordance with a previous study on human leukocyte antigen (HLA), HLA-A2 and HLA-A3 are detected with dysregulated expression patterns in colon carcinomas tissues [58]. HLA-DR disorders have also been thought to indicate colon cancer development according to Maccalli, C., *et al.*, and HLA-DRB1*1301 and HLA-DRB1*0402 were found to be associated with CRC; however, no epitopes against HLA-DRB1*1301 were found for the selected five proteins in our study [59]. Both T-cell and B-cell epitopes are able to stimulate long-acting and antigen-specific immune responses therefore have significant potential for vaccine development [81].

Overall, OMV size ranges make these ideal for vesicular transport to lymph nodes and are readily recognizable by B cells. OMVs containing moieties that have antigenic domains and epitopes are highly effective at triggering B cell activation, thereby stimulating a strong protective humoral immune response. Naturally secreted OMVs of *F. nucleatum* carry large quantities of autotransporter proteins. Thus, utilizing these OMV vehicles, which are already loaded with highly antigenic pathogen derived proteins, may facilitate the process of developing and expressing a large array of recombinant constructs for better optimized vaccines. OMVs of pathogens like *F. nucleatum* may serve as factories for the production of highly effective and molecularly designed

'ready-made' vaccines. Alternatively, OMVs can be utilized as delivery platforms loaded with one or just a few recombinant constructs proven to be highly antigenic. Our future studies will focus on the identity and verification of putative virulence factors carried by *F. nucleatum* OMVs and investigate the potential of these for vaccine design against *F. nucleatum* and colon cancer prevention. In conclusion, non-replicating OMVs from *F. nucleatum* have great potential for harnessing pathogen virulence factor antigenicity to produce an effective vaccine to combat CRC.

REFERENCE

1. Bolstad, A. I.; Jensen, H. B.; Bakken, V., Taxonomy, biology, and periodontal aspects of *Fusobacterium nucleatum*. *Clin Microbiol Rev* **1996,** *9* (1), 55-71.
2. Isomura, K., [Bacterial endotoxin from *Fusobacterium nucleatum*--particularly the role of endotoxin in periodontal diseases]. *Nihon Shishubyo Gakkai Kaishi* **1985,** *27* (2), 328-51.
3. Strauss, J.; White, A.; Ambrose, C.; McDonald, J.; Allen-Vercoe, E., Phenotypic and genotypic analyses of clinical *Fusobacterium nucleatum* and *Fusobacterium periodonticum* isolates from the human gut. *Anaerobe* **2008,** *14* (6), 301-9.
4. Kuboniwa, M.; Lamont, R. J., Subgingival biofilm formation. *Periodontol 2000* **2010,** *52* (1), 38-52.
5. Yang, N. Y.; Zhang, Q.; Li, J. L.; Yang, S. H.; Shi, Q., Progression of periodontal inflammation in adolescents is associated with increased number of *Porphyromonas gingivalis, Prevotella intermedia, Tannerella forsythensis,* and *Fusobacterium nucleatum*. *Int J Paediatr Dent* **2014,** *24* (3), 226-33.
6. Field, C. A.; Gidley, M. D.; Preshaw, P. M.; Jakubovics, N., Investigation and quantification of key periodontal pathogens in patients with type 2 diabetes. *J Periodontal Res* **2012,** *47* (4), 470-8.
7. Griffen, A. L.; Beall, C. J.; Campbell, J. H.; Firestone, N. D.; Kumar, P. S.; Yang, Z. K.; Podar, M.; Leys, E. J., Distinct and complex bacterial profiles in human periodontitis and health revealed by 16S pyrosequencing. *Isme J* **2012,** *6* (6), 1176-1185.
8. Van Palenstein Helderman, W. H., Total viable count and differential count of V*ibrio (campylobacter) sputorum, Fusobacterium nucleatum, Selenomonas sputigena, Bacteroides ochraceus* and *Veillonella* in the inflamed and non inflamed human gingival crevice. *J Periodontal Res* **1975,** *10* (5), 294-305.
9. Falkler, W. A., Jr.; Lai, R.; Vincent, J. W.; Dober, L.; Spiegel, C.; Hayduk, S., The ELISA system for measuring antibody reactive to *Fusobacterium nucleatum* in the sera of patients with chronic periodontitis. *J Periodontol* **1982,** *53* (12), 762-6.
10. Diaz, P. I.; Zilm, P. S.; Rogers, A. H., Fusobacterium nucleatum supports the growth of *Porphyromonas gingivalis* in oxygenated and carbon-dioxide-depleted environments. *Microbiology* **2002,** *148* (Pt 2), 467-72.
11. Kaufman, J.; DiRienzo, J. M., Evidence for the existence of two classes of corncob (coaggregation) receptor in Fusobacterium nucleatum. *Oral Microbiol Immunol* **1988,** *3* (4), 145-52.
12. Lancy, P., Jr.; Dirienzo, J. M.; Appelbaum, B.; Rosan, B.; Holt, S. C., Corncob formation between *Fusobacterium nucleatum* and *Streptococcus sanguis*. *Infect Immun* **1983,** *40* (1), 303-9.
13. Kaufman, J.; DiRienzo, J. M., Isolation of a corncob (coaggregation) receptor polypeptide from *Fusobacterium nucleatum*. *Infect Immun* **1989,** *57* (2), 331-7.
14. Garrett, W. S., Cancer and the microbiota. *Science* **2015,** *348* (6230), 80-86.
15. Brennan, C. A.; Garrett, W. S., Gut Microbiota, Inflammation, and Colorectal Cancer. *Annu Rev Microbiol* **2016,** *70*, 395-411.
16. Arnold, M.; Sierra, M. S.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F., Global patterns and trends in colorectal cancer incidence and mortality. *Gut* **2017,** *66* (4), 683-691.
17. Bullman, S.; Pedamallu, C. S.; Sicinska, E.; Clancy, T. E.; Zhang, X.; Cai, D.; Neuberg, D.; Huang, K.; Guevara, F.; Nelson, T.; Chipashvili, O.; Hagan, T.; Walker, M.; Ramachandran, A.; Diosdado, B.; Serna, G.; Mulet, N.; Landolfi, S.; Ramon, Y. C. S.; Fasani, R.; Aguirre, A. J.;

Ng, K.; Elez, E.; Ogino, S.; Tabernero, J.; Fuchs, C. S.; Hahn, W. C.; Nuciforo, P.; Meyerson, M., Analysis of *Fusobacterium persistence* and antibiotic response in colorectal cancer. *Science* **2017,** *358* (6369), 1443-1448.
18. Castellarin, M.; Warren, R. L.; Freeman, J. D.; Dreolini, L.; Krzywinski, M.; Strauss, J.; Barnes, R.; Watson, P.; Allen-Vercoe, E.; Moore, R. A.; Holt, R. A., *Fusobacterium nucleatum* infection is prevalent in human colorectal carcinoma. *Genome Res* **2012,** *22* (2), 299-306.
19. Flanagan, L.; Schmid, J.; Ebert, M.; Soucek, P.; Kunicka, T.; Liska, V.; Bruha, J.; Neary, P.; Dezeeuw, N.; Tommasino, M.; Jenab, M.; Prehn, J. H.; Hughes, D. J., *Fusobacterium nucleatum* associates with stages of colorectal neoplasia development, colorectal cancer and disease outcome. *Eur J Clin Microbiol Infect Dis* **2014,** *33* (8), 1381-90.
20. Kostic, A. D.; Gevers, D.; Pedamallu, C. S.; Michaud, M.; Duke, F.; Earl, A. M.; Ojesina, A. I.; Jung, J.; Bass, A. J.; Tabernero, J.; Baselga, J.; Liu, C.; Shivdasani, R. A.; Ogino, S.; Birren, B. W.; Huttenhower, C.; Garrett, W. S.; Meyerson, M., Genomic analysis identifies association of *Fusobacterium* with colorectal carcinoma. *Genome Res* **2012,** *22* (2), 292-8.
21. Gur, C.; Ibrahim, Y.; Isaacson, B.; Yamin, R.; Abed, J.; Gamliel, M.; Enk, J.; Bar-On, Y.; Stanietsky-Kaynan, N.; Coppenhagen-Glazer, S.; Shussman, N.; Almogy, G.; Cuapio, A.; Hofer, E.; Mevorach, D.; Tabib, A.; Ortenberg, R.; Markel, G.; Miklic, K.; Jonjic, S.; Brennan, C. A.; Garrett, W. S.; Bachrach, G.; Mandelboim, O., Binding of the Fap2 protein of *Fusobacterium nucleatum* to human inhibitory receptor TIGIT protects tumors from immune cell attack. *Immunity* **2015,** *42* (2), 344-355.
22. Ye, X.; Wang, R.; Bhattacharya, R.; Boulbes, D. R.; Fan, F.; Xia, L.; Adoni, H.; Ajami, N. J.; Wong, M. C.; Smith, D. P.; Petrosino, J. F.; Venable, S.; Qiao, W.; Baladandayuthapani, V.; Maru, D.; Ellis, L. M., *Fusobacterium nucleatum subspecies animalis* Influences Proinflammatory Cytokine Expression and Monocyte Activation in Human Colorectal Tumors. *Cancer Prev Res (Phila)* **2017,** *10* (7), 398-409.
23. Bettegowda, C.; Sausen, M.; Leary, R. J.; Kinde, I.; Wang, Y.; Agrawal, N.; Bartlett, B. R.; Wang, H.; Luber, B.; Alani, R. M.; Antonarakis, E. S.; Azad, N. S.; Bardelli, A.; Brem, H.; Cameron, J. L.; Lee, C. C.; Fecher, L. A.; Gallia, G. L.; Gibbs, P.; Le, D.; Giuntoli, R. L.; Goggins, M.; Hogarty, M. D.; Holdhoff, M.; Hong, S. M.; Jiao, Y.; Juhl, H. H.; Kim, J. J.; Siravegna, G.; Laheru, D. A.; Lauricella, C.; Lim, M.; Lipson, E. J.; Marie, S. K.; Netto, G. J.; Oliner, K. S.; Olivi, A.; Olsson, L.; Riggins, G. J.; Sartore-Bianchi, A.; Schmidt, K.; Shih l, M.; Oba-Shinjo, S. M.; Siena, S.; Theodorescu, D.; Tie, J.; Harkins, T. T.; Veronese, S.; Wang, T. L.; Weingart, J. D.; Wolfgang, C. L.; Wood, L. D.; Xing, D.; Hruban, R. H.; Wu, J.; Allen, P. J.; Schmidt, C. M.; Choti, M. A.; Velculescu, V. E.; Kinzler, K. W.; Vogelstein, B.; Papadopoulos, N.; Diaz, L. A., Jr., Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci Transl Med* **2014,** *6* (224), 224ra24.
24. Mima, K.; Sukawa, Y.; Nishihara, R.; Qian, Z. R.; Yamauchi, M.; Inamura, K.; Kim, S. A.; Masuda, A.; Nowak, J. A.; Nosho, K.; Kostic, A. D.; Giannakis, M.; Watanabe, H.; Bullman, S.; Milner, D. A.; Harris, C. C.; Giovannucci, E.; Garraway, L. A.; Freeman, G. J.; Dranoff, G.; Chan, A. T.; Garrett, W. S.; Huttenhower, C.; Fuchs, C. S.; Ogino, S., *Fusobacterium nucleatum* and T Cells in Colorectal Carcinoma. *JAMA Oncol* **2015,** *1* (5), 653-61.
25. Yu, T.; Guo, F.; Yu, Y.; Sun, T.; Ma, D.; Han, J.; Qian, Y.; Kryczek, I.; Sun, D.; Nagarsheth, N.; Chen, Y.; Chen, H.; Hong, J.; Zou, W.; Fang, J. Y., *Fusobacterium nucleatum* Promotes Chemoresistance to Colorectal Cancer by Modulating Autophagy. *Cell* **2017,** *170* (3), 548-563 e16.

26. Rubinstein, M. R.; Wang, X.; Liu, W.; Hao, Y.; Cai, G.; Han, Y. W., *Fusobacterium nucleatum* promotes colorectal carcinogenesis by modulating E-cadherin/beta-catenin signaling via its FadA adhesin. *Cell Host Microbe* **2013,** *14* (2), 195-206.
27. Abed, J.; Emgard, J. E.; Zamir, G.; Faroja, M.; Almogy, G.; Grenov, A.; Sol, A.; Naor, R.; Pikarsky, E.; Atlan, K. A.; Mellul, A.; Chaushu, S.; Manson, A. L.; Earl, A. M.; Ou, N.; Brennan, C. A.; Garrett, W. S.; Bachrach, G., Fap2 Mediates *Fusobacterium nucleatum* Colorectal Adenocarcinoma Enrichment by Binding to Tumor-Expressed Gal-GalNAc. *Cell Host Microbe* **2016,** *20* (2), 215-25.
28. Kostic, A. D.; Chun, E.; Robertson, L.; Glickman, J. N.; Gallini, C. A.; Michaud, M.; Clancy, T. E.; Chung, D. C.; Lochhead, P.; Hold, G. L.; El-Omar, E. M.; Brenner, D.; Fuchs, C. S.; Meyerson, M.; Garrett, W. S., *Fusobacterium nucleatum* potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. *Cell Host Microbe* **2013,** *14* (2), 207-15.
29. Xu, M.; Yamada, M.; Li, M.; Liu, H.; Chen, S. G.; Han, Y. W., FadA from *Fusobacterium nucleatum* utilizes both secreted and nonsecreted forms for functional oligomerization for attachment and invasion of host cells. *J Biol Chem* **2007,** *282* (34), 25000-9.
30. Rothwell, P. M.; Wilson, M.; Elwin, C. E.; Norrving, B.; Algra, A.; Warlow, C. P.; Meade, T. W., Long-term effect of aspirin on colorectal cancer incidence and mortality: 20-year follow-up of five randomised trials. *Lancet* **2010,** *376* (9754), 1741-50.
31. Benamouzig, R.; Uzzan, B.; Little, J.; Chaussade, S., Low dose aspirin, COX-inhibition and chemoprevention of colorectal cancer. *Curr Top Med Chem* **2005,** *5* (5), 493-503.
32. Schwechheimer, C.; Kuehn, M. J., Outer-membrane vesicles from Gram-negative bacteria: biogenesis and functions. *Nat Rev Microbiol* **2015,** *13* (10), 605-19.
33. O'Ryan, M.; Stoddard, J.; Toneatto, D.; Wassil, J.; Dull, P. M., A multi-component meningococcal serogroup B vaccine (4CMenB): the clinical development program. *Drugs* **2014,** *74* (1), 15-30.
34. Bachmann, M. F.; Jennings, G. T., Vaccine delivery: a matter of size, geometry, kinetics and molecular patterns. *Nat Rev Immunol* **2010,** *10* (11), 787-96.
35. Strauss, J.; Kaplan, G. G.; Beck, P. L.; Rioux, K.; Panaccione, R.; Devinney, R.; Lynch, T.; Allen-Vercoe, E., Invasive potential of gut mucosa-derived *Fusobacterium nucleatum* positively correlates with IBD status of the host. *Inflamm Bowel Dis* **2011,** *17* (9), 1971-8.
36. Chutkan, H.; Macdonald, I.; Manning, A.; Kuehn, M. J., Quantitative and qualitative preparations of bacterial outer membrane vesicles. *Methods Mol Biol* **2013,** *966*, 259-72.
37. Klimentova, J.; Stulik, J., Methods of isolation and purification of outer membrane vesicles from gram-negative bacteria. *Microbiol Res* **2015,** *170*, 1-9.
38. Aguilera, L.; Toloza, L.; Gimenez, R.; Odena, A.; Oliveira, E.; Aguilar, J.; Badia, J.; Baldoma, L., Proteomic analysis of outer membrane vesicles from the probiotic strain *Escherichia coli Nissle* 1917. *Proteomics* **2014,** *14* (2-3), 222-9.
39. Yang, Y.; Thannhauser, T. W.; Li, L.; Zhang, S., Development of an integrated approach for evaluation of 2-D gel image analysis: Impact of multiple proteins in single spots on comparative proteomics in conventional 2-D gel/MALDI workflow. *Electrophoresis* **2007,** *28* (12), 2080-2094.
40. Thomas, C. J.; Cleland, T. P.; Zhang, S.; Gundberg, C. M.; Vashishth, D., Identification and characterization of glycation adducts on osteocalcin. *Analytical Biochemistry* **2017,** *525*, 46-53.

41. Yang, Y.; Anderson, E.; Zhang, S., Evaluation of six sample preparation procedures for qualitative and quantitative proteomics analysis of milk fat globule membrane. *Electrophoresis* **2018**.
42. UniProt Consortium, T., UniProt: the universal protein knowledgebase. *Nucleic Acids Res* **2018,** *46* (5), 2699.
43. O'Leary, N. A.; Wright, M. W.; Brister, J. R.; Ciufo, S.; Haddad, D.; McVeigh, R.; Rajput, B.; Robbertse, B.; Smith-White, B.; Ako-Adjei, D.; Astashyn, A.; Badretdin, A.; Bao, Y.; Blinkova, O.; Brover, V.; Chetvernin, V.; Choi, J.; Cox, E.; Ermolaeva, O.; Farrell, C. M.; Goldfarb, T.; Gupta, T.; Haft, D.; Hatcher, E.; Hlavina, W.; Joardar, V. S.; Kodali, V. K.; Li, W.; Maglott, D.; Masterson, P.; McGarvey, K. M.; Murphy, M. R.; O'Neill, K.; Pujar, S.; Rangwala, S. H.; Rausch, D.; Riddick, L. D.; Schoch, C.; Shkeda, A.; Storz, S. S.; Sun, H.; Thibaud-Nissen, F.; Tolstoy, I.; Tully, R. E.; Vatsan, A. R.; Wallin, C.; Webb, D.; Wu, W.; Landrum, M. J.; Kimchi, A.; Tatusova, T.; DiCuccio, M.; Kitts, P.; Murphy, T. D.; Pruitt, K. D., Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res* **2016,** *44* (D1), D733-45.
44. Tsirigos, K. D.; Elofsson, A.; Bagos, P. G., PRED-TMBB2: improved topology prediction and detection of beta-barrel outer membrane proteins. *Bioinformatics* **2016,** *32* (17), i665-i671.
45. He, Y.; Xiang, Z.; Mobley, H. L., Vaxign: the first web-based vaccine design program for reverse vaccinology and applications for vaccine development. *J Biomed Biotechnol* **2010,** *2010*, 297505.
46. Xiang, Z. S.; He, Y. Q., Vaxign: a web-based vaccine target design program for reverse vaccinology. *Procedia Vaccinol* **2009,** *1* (1), 23-29.
47. Xiang, Z.; He, Y., Genome-wide prediction of vaccine targets for human herpes simplex viruses using Vaxign reverse vaccinology. *BMC Bioinformatics* **2013,** *14 Suppl 4*, S2.
48. Sachdeva, G.; Kumar, K.; Jain, P.; Ramachandran, S., SPAAN: a software program for prediction of adhesins and adhesin-like proteins using neural networks. *Bioinformatics* **2005,** *21* (4), 483-91.
49. Gardy, J. L.; Laird, M. R.; Chen, F.; Rey, S.; Walsh, C. J.; Ester, M.; Brinkman, F. S., PSORTb v.2.0: expanded prediction of bacterial protein subcellular localization and insights gained from comparative proteome analysis. *Bioinformatics* **2005,** *21* (5), 617-23.
50. Yu, C. S.; Cheng, C. W.; Su, W. C.; Chang, K. C.; Huang, S. W.; Hwang, J. K.; Lu, C. H., CELLO2GO: a web server for protein subCELlular LOcalization prediction with functional gene ontology annotation. *PLoS One* **2014,** *9* (6), e99368.
51. Juncker, A. S.; Willenbrock, H.; Von Heijne, G.; Brunak, S.; Nielsen, H.; Krogh, A., Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Sci* **2003,** *12* (8), 1652-62.
52. Finn, R. D.; Attwood, T. K.; Babbitt, P. C.; Bateman, A.; Bork, P.; Bridge, A. J.; Chang, H. Y.; Dosztanyi, Z.; El-Gebali, S.; Fraser, M.; Gough, J.; Haft, D.; Holliday, G. L.; Huang, H.; Huang, X.; Letunic, I.; Lopez, R.; Lu, S.; Marchler-Bauer, A.; Mi, H.; Mistry, J.; Natale, D. A.; Necci, M.; Nuka, G.; Orengo, C. A.; Park, Y.; Pesseat, S.; Piovesan, D.; Potter, S. C.; Rawlings, N. D.; Redaschi, N.; Richardson, L.; Rivoire, C.; Sangrador-Vegas, A.; Sigrist, C.; Sillitoe, I.; Smithers, B.; Squizzato, S.; Sutton, G.; Thanki, N.; Thomas, P. D.; Tosatto, S. C.; Wu, C. H.; Xenarios, I.; Yeh, L. S.; Young, S. Y.; Mitchell, A. L., InterPro in 2017-beyond protein family and domain annotations. *Nucleic Acids Res* **2017,** *45* (D1), D190-D199.

53. Bendtsen, J. D.; Kiemer, L.; Fausboll, A.; Brunak, S., Non-classical protein secretion in bacteria. *BMC Microbiol* **2005,** *5*, 58.
54. Rice, P.; Longden, I.; Bleasby, A., EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* **2000,** *16* (6), 276-7.
55. Mullan, L. J.; Bleasby, A. J., Short EMBOSS User Guide. European Molecular Biology Open Software Suite. *Brief Bioinform* **2002,** *3* (1), 92-4.
56. Olson, S. A., EMBOSS opens up sequence analysis. European Molecular Biology Open Software Suite. *Brief Bioinform* **2002,** *3* (1), 87-91.
57. Chen, J.; Liu, H.; Yang, J.; Chou, K. C., Prediction of linear B-cell epitopes using amino acid pair antigenicity scale. *Amino Acids* **2007,** *33* (3), 423-8.
58. Novellino, L.; Castelli, C.; Parmiani, G., A listing of human tumor antigens recognized by T cells: March 2004 update. *Cancer Immunol Immunother* **2005,** *54* (3), 187-207.
59. Maccalli, C.; Li, Y. F.; El-Gamil, M.; Rosenberg, S. A.; Robbins, P. F., Identification of a colorectal tumor-associated antigen (COA-1) recognized by CD4(+) T lymphocytes. *Cancer Res* **2003,** *63* (20), 6735-43.
60. Kanehisa, M.; Furumichi, M.; Tanabe, M.; Sato, Y.; Morishima, K., KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* **2017,** *45* (D1), D353-D361.
61. Huerta-Cepas, J.; Szklarczyk, D.; Forslund, K.; Cook, H.; Heller, D.; Walter, M. C.; Rattei, T.; Mende, D. R.; Sunagawa, S.; Kuhn, M.; Jensen, L. J.; von Mering, C.; Bork, P., eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res* **2016,** *44* (D1), D286-93.
62. Wickham, H., *Ggplot2 : elegant graphics for data analysis*. Springer: New York, 2009; p viii, 212 p.
63. Choi, D. S.; Kim, D. K.; Choi, S. J.; Lee, J.; Choi, J. P.; Rho, S.; Park, S. H.; Kim, Y. K.; Hwang, D.; Gho, Y. S., Proteomic analysis of outer membrane vesicles derived from Pseudomonas aeruginosa. *Proteomics* **2011,** *11* (16), 3424-9.
64. Altindis, E.; Fu, Y.; Mekalanos, J. J., Proteomic analysis of *Vibrio cholerae* outer membrane vesicles. *Proc Natl Acad Sci U S A* **2014,** *111* (15), E1548-56.
65. Veith, P. D.; Chen, Y. Y.; Gorasia, D. G.; Chen, D.; Glew, M. D.; O'Brien-Simpson, N. M.; Cecil, J. D.; Holden, J. A.; Reynolds, E. C., *Porphyromonas gingivalis* outer membrane vesicles exclusively contain outer membrane and periplasmic proteins and carry a cargo enriched with virulence factors. *J Proteome Res* **2014,** *13* (5), 2420-32.
66. Perez-Cruz, C.; Delgado, L.; Lopez-Iglesias, C.; Mercade, E., Outer-inner membrane vesicles naturally secreted by gram-negative pathogenic bacteria. *PLoS One* **2015,** *10* (1), e0116896.
67. Henderson, I. R.; Navarro-Garcia, F.; Desvaux, M.; Fernandez, R. C.; Ala'Aldeen, D., Type V protein secretion pathway: the autotransporter story. *Microbiol Mol Biol Rev* **2004,** *68* (4), 692-744.
68. Wells, T. J.; Tree, J. J.; Ulett, G. C.; Schembri, M. A., Autotransporter proteins: novel targets at the bacterial cell surface. *Fems Microbiology Letters* **2007,** *274* (2), 163-172.
69. van Ulsen, P.; Rahman, S. U.; Jong, W. S. P.; Daleke-Schermerhom, M. H.; Luirink, J., Type V secretion: From biogenesis to biotechnology. *Bba-Mol Cell Res* **2014,** *1843* (8), 1592-1611.

70. Kaplan, C. W.; Ma, X.; Paranjpe, A.; Jewett, A.; Lux, R.; Kinder-Haake, S.; Shi, W., *Fusobacterium nucleatum* outer membrane proteins Fap2 and RadD induce cell death in human lymphocytes. *Infect Immun* **2010,** *78* (11), 4773-8.
71. Henderson, I. R.; Nataro, J. P., Virulence functions of autotransporter proteins. *Infect Immun* **2001,** *69* (3), 1231-43.
72. Coppenhagen-Glazer, S.; Sol, A.; Abed, J.; Naor, R.; Zhang, X.; Han, Y. W.; Bachrach, G., Fap2 of *Fusobacterium nucleatum* is a galactose-inhibitable adhesin involved in coaggregation, cell adhesion, and preterm birth. *Infect Immun* **2015,** *83* (3), 1104-13.
73. Benz, I.; Schmidt, M. A., AIDA-I, the adhesin involved in diffuse adherence of the diarrhoeagenic *Escherichia coli* strain 2787 (O126:H27), is synthesized via a precursor molecule. *Mol Microbiol* **1992,** *6* (11), 1539-46.
74. Charbonneau, M. E.; Mourez, M., Functional organization of the autotransporter adhesin involved in diffuse adherence. *J Bacteriol* **2007,** *189* (24), 9020-9.
75. Benz, I.; Schmidt, M. A., Structures and functions of autotransporter proteins in microbial pathogens. *Int J Med Microbiol* **2011,** *301* (6), 461-8.
76. Suhr, M.; Benz, I.; Schmidt, M. A., Processing of the AIDA-I precursor: removal of AIDAc and evidence for the outer membrane anchoring as a beta-barrel structure. *Mol Microbiol* **1996,** *22* (1), 31-42.
77. Daleke-Schermerhorn, M. H.; Felix, T.; Soprova, Z.; Ten Hagen-Jongman, C. M.; Vikstrom, D.; Majlessi, L.; Beskers, J.; Follmann, F.; de Punder, K.; van der Wel, N. N.; Baumgarten, T.; Pham, T. V.; Piersma, S. R.; Jimenez, C. R.; van Ulsen, P.; de Gier, J. W.; Leclerc, C.; Jong, W. S.; Luirink, J., Decoration of outer membrane vesicles with multiple antigens by using an autotransporter approach. *Appl Environ Microbiol* **2014,** *80* (18), 5854-65.
78. Fardini, Y.; Wang, X.; Temoin, S.; Nithianantham, S.; Lee, D.; Shoham, M.; Han, Y. W., Fusobacterium nucleatum adhesin FadA binds vascular endothelial cadherin and alters endothelial integrity. *Mol Microbiol* **2011,** *82* (6), 1468-80.
79. Manson McGuire, A.; Cochrane, K.; Griggs, A. D.; Haas, B. J.; Abeel, T.; Zeng, Q.; Nice, J. B.; MacDonald, H.; Birren, B. W.; Berger, B. W.; Allen-Vercoe, E.; Earl, A. M., Evolution of invasion in a diverse set of *Fusobacterium* species. *MBio* **2014,** *5* (6), e01864.
80. Heise, T.; Dersch, P., Identification of a domain in *Yersinia* virulence factor YadA that is crucial for extracellular matrix-specific cell adhesion and uptake. *P Natl Acad Sci USA* **2006,** *103* (9), 3375-3380.
81. Ellis, R. W., New technologies for making vaccines. *Vaccine* **1999,** *17* (13-14), 1596-604.

CHAPTER THREE

CONCLUSION

During the past several years, investigations into gut microbe role in human diseases, particular in carcinogenesis, has gained much attention. Detailed reviews of metagenomic data sets comparing healthy and diseased individuals have implicated invasive gut microbial lines in the onset or progression of diseases. *F. nucleatum* has recently emerged as being associated with, if not the causative agent of, colorectal cancer (CRC); however, mechanisms of *F. nucleatum* virulence factors in promoting tumorigenesis have only begun to be revealed. OMVs secreted by *F. nucleatum* and other Gram-negative bacteria participate multiple pathways to facilitate bacterial survival in the host. OMVs are highly immunogenic and have recently been developed as novel platforms for safe and efficient delivery of vaccines. Overall, the OMV size ranges make them ideal for vesicular transport to lymph nodes and are readily recognized by B cells. OMVs containing moieties that have antigenic domains and epitopes are highly effective at triggering B cell activation, thereby stimulating a strong protective antibody response. Naturally secreted OMVs of *F. nucleatum* carry large quantities of autotransporter proteins; utilizing these vehicles, which are already loaded with highly antigenic pathogen derived proteins may allow researchers to mitigate, if not to avoid, the labor-intensive process of developing and expressing a large array of recombinant constructs. In other words, cultures of pathogens like *F. nucleatum* may serve as factories for the production of highly effective and specific 'ready-made' vaccines. Alternatively, OMVs can be utilized as delivery platforms loaded with one or just a few recombinant constructs proven to be highly antigenic.

Our future studies will focus on the identification and verification of putative virulence factors carried by *F. nucleatum* OMVs and investigate the potential for these in vaccine designs against colon cancer. The protective ability of *F. nucleatum* OMVs will be tested by inoculating into mice and evaluating the immune responses. Selected proteins with antigenic potential will be cloned and expressed as recombinant proteins (with C-terminal His6-tag for protein purification). Individual purified proteins will be injected into mice and monitored for immune responses. Further assays will be used to identify the protective activity of putative vaccine candidates. Additionally, specific protein mutant strains will also be constructed to explore target protein function.

We have characterized the proteome of OMVs from *F. nucleatum* and provide topological analysis of proteins these contain; OMVs have been found to be enriched with autotransporter proteins. This suggests that OMVs may represent important cargo carrying functions and provide specific pathways for virulence and nutrient resources. In conclusion, non-replicating OMVs from *F. nucleatum* have great potential for harnessing pathogen virulence factor antigenicity to produce an effective vaccine to combat CRC. The abundant autotransporter proteins carried by OMVs are expected to elicit sufficient immune responses. Moreover, numerous antigenic epitopes exposed on nanoparticles are designed to trigger robust B cells and T cell responses.